Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/EP2003/013655

International filing date: 03 December 2003 (03.12.2003)

Document type: Certified copy of priority document

Document details: Country/Office: EP
Number: 02 026 921.3
Filing date: 03 December 2002 (03.12.2002)

Date of receipt at the International Bureau: 22 December 2006 (22.12.2006)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)

World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

EP/03/13655

Europäisches Patentamt | European Patent Office | Office européen des brevets

REC'D 22 DEC 2006
WIPO PCT

Bescheinigung | Certificate | Attestation

Die angehefteten Unterlagen stimmen mit der ursprünglich eingereichten Fassung der auf dem nächsten Blatt bezeichneten europäischen Patentanmeldung überein.

The attached documents are exact copies of the European patent application described on the following page, as originally filed.

Les documents fixés à cette attestation sont conformes à la version initialement déposée de la demande de brevet européen spécifiée à la page suivante.

Patentanmeldung Nr. | Patent application No. | Demande de brevet n°

02026921.3

Der Präsident des Europäischen Patentamts;
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets
p.o.

R C van Dijk

Anmeldung Nr:
Application no.: 02026921.3
Demande no:

Anmeldetag:
Date of filing: 03.12.02
Date de dépôt:

Anmelder/Applicant(s)/Demandeur(s):

DeveloGen Aktiengesellschaft für
entwicklungsbiologische Forschung
Rudolf-Wissell-Strasse 28
37079 Göttingen
ALLEMAGNE

Bezeichnung der Erfindung/Title of the invention/Titre de l'invention:
(Falls die Bezeichnung der Erfindung nicht angegeben ist, siehe Beschreibung.
If no title is shown please refer to the description.
Si aucun titre n'est indiqué se referer à la description.)

Prl homologous proteins involved in the regulation of energy homeostasis

In Anspruch genommene Priorität(en) / Priority(ies) claimed /Priorité(s) revendiquée(s)
Staat/Tag/Aktenzeichen/State/Date/File no./Pays/Date/Numéro de dépôt:

Internationale Patentklassifikation/International Patent Classification/
Classification internationale des brevets:

A61K38/00

Am Anmeldetag benannte Vertragstaaten/Contracting states designated at date of filing/Etats contractants désignées lors du dépôt:

AT BE BG CH CY CZ DE DK EE ES FI FR GB GR IE IT LI LU MC NL PT SE SI SK

03. Dez. 2002

WEICKMANN & WEICKMANN

Patentanwälte
European Patent Attorneys · European Trademark Attorneys

EPO - Munich
69
03. Dez. 2002

DIPL.-ING. H. WEICKMANN (bis 31.1.01)
DIPL.-ING. F. A. WEICKMANN
DIPL.-CHEM. B. HUBER
DR.-ING. H. LISKA
DIPL.-PHYS. DR. J. PRECHTEL
DIPL.-CHEM. DR. B. BÖHM
DIPL.-CHEM. DR. W. WEISS
DIPL.-PHYS. DR. J. TIESMEYER
DIPL.-PHYS. DR. M. HERZOG
DIPL.-PHYS. B. RUTTENSPERGER
DIPL.-PHYS. DR.-ING. V. JORDAN
DIPL.-CHEM. DR. M. DEY
DIPL.-FORSTW. DR. J. LACHNIT

Unser Zeichen:
29518P EP/WWld

Anmelder:
DeveloGen Aktiengesellschaft für
entwicklungsbiologische Forschung
Rudolf-Wissell-Straße 28

37079 Göttingen

Prl homologous proteins involved in the regulation of energy homeostasis

---

Postfach 860 820, 81635 München, Deutschland, Tel. (089) 45563 0, Fax (089) 45563 999, email@weickmann.de

EPO - Munich
69
03. Dez. 2002

Prl homologous proteins involved in the regulation of energy homeostasis

Description

This invention relates to the use of nucleic acid sequences encoding mammalian protein tyrosine phosphatase (referred to as Prl-1), and the polypeptides encoded thereby and to the use thereof or effectors of Prl-1 in the diagnosis, study, prevention, and treatment of diseases and disorders related to body-weight regulation, for example, but not limited to, metabolic diseases or dysfunctions such as obesity as well as related disorders such as Type 2 diabetes.

There are several metabolic diseases of human and animal metabolism, eg., obesity and severe weight loss, that relate to energy imbalance where caloric intake versus energy expenditure is imbalanced. Obesity is one of the most prevalent metabolic disorders in the world. It is still a poorly understood human disease that becomes as a major health problem more and more relevant for western society. Obesity is defined as a body weight more than 20% in excess of the ideal body weight, frequently resulting in a significant impairment of health. Obesity may be measured by body mass index, an indicator of adiposity or fatness. Further parameters for defining obesity are waist circumferences, skinfold thickness and bioimpedance (see, inter alia, Kopelman (1999), loc. cit.). It is associated with an increased risk for cardiovascular disease, hypertension, diabetes mellitus Type II, hyperlipidaemia and an increased mortality rate. Besides severe risks of illness, individuals suffering from obesity are often isolated socially.

Obesity is influenced by genetic, metabolic, biochemical, psychological, and behavioral factors and can be caused by different reasons such as

non-insulin dependent diabetes, increase in triglycerides, increase in carbohydrate bound energy and low energy expenditure. As such, it is a complex disorder that must be addressed on several fronts to achieve lasting positive clinical outcome. Since obesity is not to be considered as a single disorder but as a heterogeneous group of conditions with (potential) multiple causes, it is also characterized by elevated fasting plasma insulin and an exaggerated insulin response to oral glucose intake (Koltermann, J. Clin. Invest 65, 1980, 1272-1284). A clear involvement of obesity in type 2 diabetes mellitus can be confirmed (Kopelman, Nature 404, 2000, 635-643).

Hyperlipidemia and elevation of free fatty acids correlate clearly with the Metabolic Syndrome', which is defined as the linkage between several diseases, including obesity an insulin resistance. This often occurs in the same patients and is a major risk factor for development of Type 2 diabetes and cardiovascular disease. It was suggested that the control of lipid levels and glucose levels is required to treat Type 2 Diabetes, heart disease, and other occurances of Metabolic Syndrome (see, for example, Santomauro et al., Diabetes, 48:2836-2841).

The molecular factors regulating food intake and body weight balance are incompletely understood. Even if several candidate genes have been described which are supposed to influence the homeostatic system(s) that regulate body mass/weight, like leptin, VCPI, VCPL or the peroxisome proliferator-activated receptor-gamma co-activator, the distinct molecular mechanisms and/or molecules influencing obesity or body weight/body mass regulations are not known. In addition, several single-gene mutations resulting in obesity have been described in mice, implicating genetic factors in the etiology of obesity. (Friedman and Leibel, 1990, Cell 69: 217-220). In the ob mouse a single gene mutation (obese) results in profound obesity, which is accompanied by diabetes (Friedman et. al., 1991, Genomics 11: 1054-1062).

Therefore, the technical problem underlying the present invention was to provide for means and methods for modulating (pathological) metabolic conditions influencing body-weight regulation and/or energy homeostatic circuits. The solution to said technical problem is achieved by providing the embodiments characterized in the claims.

Accordingly, the present invention relates to genes with novel functions in body-weight regulation, energy homeostasis, metabolism, and obesity. The present invention discloses specific genes involved in the regulation of body-weight, energy homeostasis, metabolism, and obesity, as well as related diseases such as diabetes mellitus, eating disorder, cachexia, hypertension, coronary heart disease, hypercholesterolemia (dyslipidemia), and gallstones. In particular, the present invention describes the human Prl-1 homologous genes as being involved in those conditions mentioned above.

Reversible protein phosphorylation is the main strategy for controlling activities of eukaryotic cells. Protein tyrosine phosphatases (PTPs) form a large family of enzymes that serve as key regulatory components in signal transduction pathways by removing phosphate groups from proteins. Inappropriate regulation of PTP activity leads to wrong levels of tyrosine phosphorylation, which contributes to the development of many human diseases. The protein tyrosine phosphatases (PTPs), which hydrolyze the phosphate monoesters of tyrosine residues, all share a common active site motif and are classified into 3 groups. These include the receptor-like PTPs, the intracellular PTPs, and the dual-specificity PTPs, which can dephosphorylate at serine and threonine residues as well as at tyrosines. As a member of a fourth class a PTP from regenerating rat liver was described by Diamond et al. (Diamont et al., 1994, Mol. Cell Biol. 14:3752-3762). The gene, which they designated Prl-1, was one of many immediate-early genes and expressed mainly in the nucleus. Prl-1

dephosporylates ATF-7 in vitro, which is a transcription factor regulating UCP-1 expression (Peters et al., 2001,. Biol Chem 276(17):13718-26). By using an in vitro prenylation screen, Castes et al. isolated 2 human cDNAs encoding Prl-1 homologs, designated PTP(CAAX1) and PTP(CAAX2) (Prl-2), that are farnesylated in vitro by mammalian farnesyl protein transferase. Overexpression of these PTPs in epithelial cells caused a transformed phenotype in cultured cells and tumor growth in nude mice (Cates et al., 1996, Cancer Lett 110(1-2):49-55). Matter et al. identified a cDNA encoding PTP4A3, which they termed Prl-3. The deduced Prl-3 protein is 76% identical to Prl-1 (PTP4A1; 601585) and 96% identical to mouse Prl-3 (Matter et. al., 2001,Biochem Biophys Res Commun 25;283(5):1061-8). Northern blot analysis revealed expression of an approximately 2.3-kb Prl-3 transcript predominantly in heart and skeletal muscle, with lower expression in pancreas.

So far, it has not been described that the proteins of the invention and homologous proteins are involved in the regulation of energy homeostasis and body-weight regulation and related disorders, and thus, no functions in metabolic diseases and dysfunctions and other diseases as listed above have been discussed.

In this invention we refer to the proteins encoded by Drosophila Prl-1 gene and homologous orthologs, preferably human and mice, homologous polypeptides or proteins or sequences encoding those proteins as proteins of the invention.

The present invention discloses that Prl-1 homologous proteins are regulating the energy homeostasis and fat metabolism especially the metabolism and storage of triglycerides, and polynucleotides, which identify and encode the proteins disclosed in this invention. The invention also relates to vectors, host cells, antibodies, and recombinant methods for producing the polypeptides and polynucleotides of the invention. The

invention also relates to the use of these polynucleotides, polypeptides and effectors thereof in the diagnosis, study, prevention, and treatment of metabolic diseases or dysfunctions, for example, but not limited to, metabolic syndrome including obesity, diabetes mellitus, eating disorder, cachexia, hypertension, coronary heart disease, hypercholesterolemia (dyslipidemia), and gallstones.

Prl-1 homologous proteins and nucleic acid molecules coding therefore are obtainable from insect or vertebrate species, e.g. mammals or birds. Particularly preferred are nucleic acids encoding the human Prl-1 homologs (in particular the human Prl-1, Prl-2, and Prl-3 isoforms).

The invention particularly relates to a nucleic acid molecule encoding a polypeptide contributing to regulating the energy homeostasis and the metabolism of triglycerides, wherein said nucleic acid molecule comprises

(a) the nucleotide sequence of Drosophila Prl-1, human Prl-1 homologs (in particular the human Prl-1, Prl-2, and Prl-3 isoforms), and/or a sequence complementary thereto,

(b) a nucleotide sequence which hybridizes at 50°C in a solution containing 1 x SSC and 0.1% SDS to a sequence of (a),

(c) a sequence corresponding to the sequences of (a) or (b) within the degeneration of the genetic code,

(d) a sequence which encodes a polypeptide which is at least 85%, preferably at least 90%, more preferably at least 95%, more preferably at least 98% and up to 99,6% identical to the amino acid sequences of the Prl-1 protein, preferably of the human Prl-1 homologs (in particular the human Prl-1, Prl-2, and Prl-3 isoforms),

(e) a sequence which differs from the nucleic acid molecule of (a) to (d) by mutation and wherein said mutation causes an alteration, deletion, duplication and/or premature stop in the encoded polypeptide or

(f) a partial sequence of any of the nucleotide sequences of (a) to (e) having a length of at least 15 bases, preferably at least 20 bases, more preferably at least 25 bases and most preferably at least 50 bases.

The present invention relates to genes with novel functions in body-weight regulation, energy homeostasis, metabolism, and obesity, fragments of said genes, polypeptides encoded by said genes or fragments thereof, and effectors e.g. antibodies, biologically active nucleic acids, such as antisense molecules, RNAi molecules or ribozymes, aptamers, peptides or low-molecular weight organic compounds recognizing said polynucleotides or polypeptides.

The ability to manipulate and screen the genomes of model organisms such as the fly Drosophila melanogaster provides a powerful tool to analyze biological and biochemical processes that have direct relevance to more complex vertebrate organisms due to significant evolutionary conservation of genes, cellular processes, and pathways (see, for example, Adams et al., (2000) Science 287: 2185-2195). Identification of novel gene functions in model organisms can directly contribute to the elucidation of correlative pathways in mammals (humans) and of methods of modulating them. A correlation between a pathology model (such as changes in triglyceride levels as indication for metabolic syndrome including obesity) and the modified expression of a fly gene can identify the association of the human ortholog with the particular human disease.

In one embodiment, a forward genetic screen is performed in fly displaying a mutant phenotype due to misexpression of a known gene (see, Johnston Nat Rev Genet 3: 176-188 (2002); Rorth, (1996) Proc Natl Acad Sci U S A 93: 12418-12422). In this invention, we have used a genetic screen to identify mutations of Prl-1 homologous genes that cause changes in the body weight which is reflected by a significant change of triglyceride

levels. Triglycerides levels reflect the status of energy storage in cells and are significantly increased in obese patients.

One resource for screening was a Drosophila melanogaster stock collection of EP-lines. The P-vector of this collection has Gal4-UAS-binding sites fused to a basal promoter that can transcribe adjacent genomic Drosophila sequences upon binding of Gal4 to UAS-sites (Brand & Perrimon (1993) Development 118:401-415; Rorth P., (1996) Proc Natl Acad Sci U S A 93:12418-12422). This enables the EP-line collection for overexpression of endogenous flanking gene sequences. In addition, without activation of the UAS-sites, integration of the EP-element into the gene is likely to cause a reduction of gene activity, and allows determining its function by evaluating the loss-of-function phenotype.

To isolate genes with a function in energy homeostasis, several thousand EP-lines were tested for their triglyceride content after a prolonged feeding period (see Examples for more detail). Lines with significantly changed triglyceride content were selected as positive candidates for further analysis. The change of triglyceride content due to the loss of a gene function suggests gene activities in energy homeostasis in a dose dependent manner that control the amount of energy stored as triglycerides.

In this invention, the content of triglycerides of a pool of flies with the same genotype was analyzed after feeding for six days using a triglyceride assay. Male flies homozygous for the integration of vectors for Drosophila line HD-EP(2)20261 were analyzed in assays measuring the triglyceride and glycogen contents of these flies, illustrated in more detail in the EXAMPLES section. The results of the triglyceride and glycogen content analysis are shown in FIGURE 1.

Genomic DNA sequences were isolated that are localized to the EP vector (herein HD-EP(2)20261) integration. Using those isolated genomic

sequences public databases like Berkeley Drosophila Genome Project (GadFly; see also FlyBase (1999) Nucleic Acids Research 27:85-88) were screened thereby identifying the integration site of the vectors, and the corresponding gene, described in more detail in the EXAMPLES section. The molecular organization of the gene is shown in FIGURE 2.

In one embodiment of the invention, we clearly show that mammalian Prl-1 (or variants thereof) has a function in regulating the metabolism of mature adipocytes. We could show in this invention that Prl-1 is expressed in varieties of mammalian (mouse) tissues, with highest levels of expression in metabolic active tissue such as brown adipose tissue (BAT) (see Figure 4A). Further, we show that the mammalian homologue of the Drosophilia Prl-1 gene is regulated in mice with genetically induced obesity. We used mouse models of insulin resistance and/or diabetes, such as mice carrying gene knockouts in the leptin pathway (for example, ob (leptin) or db (leptin receptor/ligand) mice) to study the expression of the protein of the invention. Such mice develop typical symptoms of diabetes, show hepatic lipid accumulation and frequently have increased plasma lipid levels (see Bruning et al, 1998, Mol. Cell. 2:449-569). We found, for example, that the expression of Prl-1 is strongly upregulated in metabolic active tissue BAT compared to white adipose tissue (WAT) in genetically induced obese mice (ob-ob and db-db) (Figure 4B).

Susceptible wild type mice that show symptoms of diabetes, lipid accumulation, and high plasma lipid levels, on high fat diet (HFD), shown significantly upregulated expression of Prl-1 in brown adipose tissue, supporting a hypothesis that Prl-1 is a modulator of adipogenesis. In addition, Prl-1 is upregulated in fasted mice (Figure 4B).

In another embodiment, using a model of adipocyte differentiation, we found that in Prl-1 overexpressing cells triglyceride levels were significantly increased on day 12 of differentiation (see Figure 5A). This increase is also

seen in control experiments with known regulators of adipogenesis, such as for example PPAR gamma-1.

Glycogen levels in cells are more variable than triglyceride levels because the turnover of glycogen is higher. Glucose is taken up by the cells rapidly and stored in the form of glycogen. This energy storage is then used as a first quick response to the metabolic demands of the cell. On day 12 of differentiation of adipocytes, intracellular glycogen levels are significantly increased, indicating that the Prl-1 effects central metabolic pathways in the cell (see Figure 5B).

The invention also encompasses polynucleotides that encode the proteins of the invention and homologous proteins. Accordingly, any nucleic acid sequence, which encodes the amino acid sequences of the proteins of the invention and homologous proteins, can be used to generate recombinant molecules that express the proteins of the invention and homologous proteins. In a particular embodiment, the invention encompasses a nucleic acid encoding Drosophila Prl-1 or human Prl-1 homologs; referred to herein as the proteins of the invention. It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of nucleotide sequences encoding the proteins, some bearing minimal homology to the nucleotide sequences of any known and naturally occurring gene, may be produced. The invention contemplates each and every possible variation of nucleotide sequence that can be made by selecting combinations based on possible codon choices.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed nucleotide sequences, and in particular, those of the polynucleotide encoding the proteins of the invention, under various conditions of stringency. Hybridization conditions are based on the melting temperature (Tm) of the nucleic acid binding complex or probe, as taught in Wahl & Berger (1987: Methods Enzymol.

152:399-407) and Kimmel (1987; Methods Enzymol. 152:507-511), and may be used at a defined stringency. Preferably, hybridization under stringent conditions means that after washing for 1 h with 1 x SSC and 0.1% SDS at 50°C, preferably at 55°C, more preferably at 62°C, and most preferably at 65°C, particularly for 1 h in 0.2 x SSC and 0.1% SDS at 50°C, preferably at 55°C, more preferably at 62°C, and most preferably 65°C a positive hybridization signal is observed. Altered nucleic acid sequences encoding the proteins which are encompassed by the invention include deletions, insertions or substitutions of different nucleotides resulting in a polynucleotide that encodes the same or a functionally equivalent protein.

The encoded proteins may also contain deletions, insertions or substitutions of amino acid residues, which produce a silent change and result in functionally equivalent proteins. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues as long as the biological activity of the protein is retained. Furthermore, the invention relates to peptide fragments of the proteins or derivatives thereof such as cyclic peptides, retro-inverso peptides or peptide mimetics having a length of at least 4, preferably at least 6 and up to 50 amino acids.

Also included within the scope of the present invention are alleles of the genes encoding the proteins of the invention and homologous proteins. As used herein, an 'allele' or 'allelic sequence' is an alternative form of the gene, which may result from at least one mutation in the nucleic acid sequence. Alleles may result in altered mRNAs or polypeptides whose structures or function may or may not be altered. Any given gene may have none, one or many allelic forms. Common mutational changes, which give rise to alleles, are generally ascribed to natural deletions, additions or substitutions of nucleotides. Each of these types of changes may occur

alone or in combination with the others, one or more times in a given sequence.

The nucleic acid sequences encoding the proteins of the invention and homologous proteins may be extended utilizing a partial nucleotide sequence and employing various methods known in the art to detect upstream sequences such as promoters and regulatory elements.

In order to express a biologically active protein, the nucleotide sequences encoding the proteins or functional equivalents, may be inserted into appropriate expression vectors, i.e., a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods, which are well known to those skilled in the art, may be used to construct expression vectors containing sequences encoding the proteins and the appropriate transcriptional and translational control elements. Regulatory elements include for example a promoter, an initiation codon, a stop codon, a mRNA stability regulatory element, and a polyadenylation signal. Expression of a polynucleotide can be assured by (i) constitutive promoters such as the Cytomegalovirus (CMV) promoter/enhancer region, (ii) tissue specific promoters such as the insulin promoter (see, Soria et al., 2000, Diabetes 49:157), SOX2 gene promotor (see Li et al., 1998, Curr. Biol. 8:971-4), Msi-1 promotor (see Sakakibara et al., 1997, J. Neuroscience 17:8300-8312), alpha-cardia myosin heavy chain promotor or human atrial natriuretic factor promotor (Klug et al., 1996, J. clin. Invest 98:216-24; Wu et al., 1989, J. Biol. Chem. 264:6472-79) or (iii) inducible promoters such as the tetracycline inducible system. Expression vectors can also contain a selection agent or marker gene that confers antibiotic resistance such as the neomycin, hygromycin or puromycin resistance genes. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. Such techniques are described in Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview, N.Y.

and Ausubel, F.M. et al. (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York, N.Y.

In a further embodiment of the invention, natural, modified or recombinant nucleic acid sequences encoding the proteins of the invention and homologous proteins may be ligated to a heterologous sequence to encode a fusion protein.

A variety of expression vector/host systems may be utilized to contain and express sequences encoding the proteins or fusion proteins. These include, but are not limited to, micro-organisms such as bacteria transformed with recombinant bacteriophage, plasmid or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (e.g., baculovirus, adenovirus, adeno-associated virus, lentiverus, retrovirus); plant cell systems transformed with virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or PBR322 plasmids); or animal cell systems.

The presence of polynucleotide sequences of the invention in a sample can be detected by DNA-DNA or DNA-RNA hybridization and/or amplification using probes or portions or fragments of said polynucleotides. Nucleic acid amplification based assays involve the use of oligonucleotides or oligomers based on the sequences specific for the gene to detect transformants containing DNA or RNA encoding the corresponding protein. As used herein 'oligonucleotides' or 'oligomers' refer to a nucleic acid sequence of at least about 10 nucleotides and as many as about 60 nucleotides, preferably about 15 to 30 nucleotides, and more preferably about 20-25 nucleotides, which can be used as a probe or amplimer.

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid

assays. Means for producing labeled hybridization or PCR probes for detecting polynucleotide seqüences include oligo-labeling, nick translation, end-labeling of labeled RNA probes, PCR amplification using a labeled nucleotide, or enzymatic synthesis. These procedures may be conducted using a variety of commercially available kits (Pharmacia & Upjohn, (Kalamazoo, Mich.); Promega (Madison Wis.); and U.S. Biochemical Corp., (Cleveland, Ohio).

The presence of proteins of the invention in a sample can be determined by immunological methods or activity measurement. A variety of protocols for detecting and measuring the expression of proteins, using either polyclonal or monoclonal antibodies specific for the protein or reagents for determining protein activity are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on the protein is preferred, but a competitive binding assay may be employed. These and other assays are described, among other places, in Hampton, R. et al. (1990; Serological Methods, a Laboratory Manual, APS Press, St Paul, Minn.) and Maddox, D. E. et al. (1983; J. Exp. Med. 158:1211-1216).

Suitable reporter molecules or labels, which may be used, include radionuclides, enzymes, fluorescent, chemiluminescent or chromogenic agents as well as substrates, co-factors, inhibitors, magnetic particles, and the like.

The nucleic acids encoding the proteins of the invention can be used to generate transgenic animal or site specific gene modifications in cell lines. Transgenic animals may be made through homologous recombination, where the normal locus of the genes encoding the proteins of the invention is altered. Alternatively, a nucleic acid construct is randomly integrated into

the genome. Vectors for stable integration include plasmids, retrovirusses and other animal virusses, YACs, and the like. The modified cells or animal are useful in the study of the function and regulation of the proteins of the invention. For example, a series of small deletions and/or substitutions may be made in the genes that encode the proteins of the invention to determine the role of particular domains of the protein, functions in pancreatic differentiation, etc.

Specific constructs of interest include anti-sense molecules, which will block the expression of the proteins of the invention, or expression of dominant negative mutations. A detectable marker, such as for example lac-Z, may be introduced in the locus of the genes of the invention, where upregulation of expression of the genes of the invention will result in an easily detected change in phenotype.

One may also provide for expression of the genes of the invention or variants thereof in cells or tissues where it is not normally expressed or at abnormal times of development. In addition, by providing expression of the proteins of the invention in cells in which they are not normally produced, one can induce changes in cell behavior.

DNA constructs for homologous recombination will comprise at least portions of the genes of the invention with the desired genetic modification, and will include regions of homology to the target locus. DNA constructs for random integration need not include regions of homology to mediate recombination. Conveniently, markers for positive and/or negative selection are included. Methods for generating cells having targeted gene modifications through homologous recombination are known in the art. For embryonic stem (ES) cells, an ES cell line may be employed, or embryonic cells may be obtained freshly from a host, e.g. mouse, rat, guinea pig etc. Such cells are grown on an appropriate fibroblast-feeder layer or grown in presence of leukemia inhibiting factor (LIF).

When ES or embryonic cells or somatic pluripotent stem cells have been transformed, they may be used to produce transgenic animals. After transformation, the cells are plated onto a feeder layer in an appropriate medium. Cells containing the construct may be detected by employing a selective medium. After sufficient time for colonies to grow, they are picked and analyzed for the occurrence of homologous recombination or integration of the construct. Those colonies that are positive may then be used for embryo manipulation and blastocyst injection. Blastocysts are obtained from 4 to 6 week old superovulated females. The ES cells are trypsinized, and the modified cells are injected into the blastocoel of the blastocyst. After injection, the blastocysts are returned to each uterine horn of pseudopregnant females. Females are then allowed to go to term and the resulting offspring screened for the construct. By providing for a different phenotype of the blastocyst and the genetically modified cells, chimeric progeny can be readily detected. The chimeric animals are screened for the presence of the modified gene and males and females having the modification are mated to produce homozygous progeny. If the gene alterations cause lethality at some point in development, tissues or organs can be maintained as allogenic or congenic grafts or transplants, or in vitro culture. The transgenic animals may be any non-human mammal, such as laboratory animal, domestic animals, etc. The transgenic animals may be used in functional studies, drug screening, etc.

Diagnostics and Therapeutics

The data disclosed in this invention show that the nucleic acids and proteins of the invention and effector molecules thereof are useful in diagnostic and therapeutic applications implicated, for example, but not limited to, metabolic syndrome including obesity, diabetes mellitus, eating disorder, cachexia, hypertension, coronary heart disease, hypercholesterolemia (dyslipidemia), and gallstones.. Hence, diagnostic and

therapeutic uses for the proteins of the invention nucleic acids and proteins of the invention are, for example but not limited to, the following: (i) protein therapeutic, (ii) small molecule drug target, (iii) antibody target (therapeutic, diagnostic, drug targeting/cytotoxic antibody), (iv) diagnostic and/or prognostic marker, (v) gene therapy (gene delivery/gene ablation), (vi) research tools, and (vii) tissue regeneration in vitro and in vivo (regeneration for all these tissues and cell types composing these tissues and cell types derived from these tissues).

The nucleic acids and proteins of the invention and effectors thereof are useful in diagnostic and therapeutic applications implicated in various applications as described below. For example, but not limited to, cDNAs encoding the proteins of the invention and particularly their human homologues may be useful in gene therapy, and the proteins of the invention and particularly their human homologues may be useful when administered to a subject in need thereof. By way of non-limiting example, the compositions of the present invention will have efficacy for treatment of patients suffering from, for example, but not limited to, in metabolic disorders as described above.

The nucleic acids of the invention or fragments thereof, may further be useful in diagnostic applications, wherein the presence or amount of the nucleic acids or the proteins are to be assessed. Further antibodies that bind immunospecifically to the substances of the invention may be used in therapeutic or diagnostic methods.

For example, in one aspect, antibodies, which are specific for the proteins of the invention and homologous proteins, may be used directly as an effector, e.g. an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissue which express the protein. The antibodies may be generated using methods that are well known in the art. Such antibodies may include, but are not limited

to, polyclonal, monoclonal, chimeric single chain, Fab fragments, and fragments produced by a Fab expression library. Neutralising antibodies, (i.e., those which inhibit dimer formation) are especially preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others, may be immunized by injection with the protein or any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. It is preferred that the peptides, fragments or oligopeptides used to induce antibodies to the protein have an amino acid sequence consisting of at least five amino acids, and more preferably at least 10 amino acids.

Monoclonal antibodies to the proteins may be prepared using any technique that provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Köhler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R. J. et al. Proc. Natl. Acad. Sci. 80:2026-2030; Cole, S. P. et al. (1984) Mol. Cell Biol. 62:109-120).

In addition, techniques developed for the production of 'chimeric antibodies', the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity can be used (Morrison, S. L. et al. (1984) Proc. Natl. Acad. Sci. 81:6851-6855; Neuberger, M. S. et al (1984) Nature 312:604-608; Takeda, S. et al. (1985) Nature 314:452-454). Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce single chain antibodies specific for the proteins of the invention and homologous proteins. Antibodies with

related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries (Burton, D. R. (1991) Proc. Natl. Acad. Sci. 88:11120-3). Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening recombinant immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature (Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299).

Antibody fragments which contain specific binding sites for the proteins may also be generated. For example, such fragments include, but are not limited to, the $F(ab')_2$ fragments which can be produced by Pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of $F(ab')_2$ fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity (Huse, W. D. et al. (1989) Science 254:1275-1281).

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding and immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between the protein and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reacive to two non-interfering protein epitopes are preferred, but a competitive binding assay may also be employed (Maddox, supra).

In another embodiment of the invention, the polynucleotides of the invention or fragments thereof or nucleic acid effector molecules such as aptamers, antisense molecules, RNAi molecules or ribozymes may be used for therapeutic purposes. In one aspect, aptamers, i.e. nucleic acid

molecules, which are capable of binding to a Prl protein and modulating its activity, may be generated by a screening and selection procedure involving the use of combinatorial nucleic acid libraries.

In a further aspect, antisense molecules may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding the proteins of the invention and homologous proteins. Thus, antisense molecules may be used to modulate protein activity or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligomers or larger fragments, can be designed from various locations along the coding or control regions of sequences encoding the proteins. Expression vectors derived from retroviruses, adenovirus, herpes or vaccinia viruses or from various bacterial plasmids may be used for delivery of nucleotide sequences to the targeted organ, tissue or cell population. Methods, which are well known to those skilled in the art, can be used to construct recombinant vectors, which will express antisense molecules complementary to the polynucleotides of the genes encoding the proteins of the invention and homologous proteins. These techniques are described both in Sambrook et al. (supra) and in Ausubel et al. (supra). Genes encoding the proteins of the invention and homologous proteins can be turned off by transforming a cell or tissue with expression vectors, which express high levels of polynucleotides that encode the proteins of the invention and homologous proteins or fragments thereof. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector and even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing antisense molecules, e.g. DNA, RNA or nucleic acid analogues such as PNA, to the control regions of the genes encoding the proteins of the invention and homologous proteins, i.e., the promoters, enhancers, and introns. Oligonucleotides derived from the transcription initiation site, e.g., between positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using "triple helix" base-pairing methodology. Triple helix pairing is useful because it cause inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature (Gee, J. E. et al. (1994) In; Huber, B. E. and B. I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, N.Y.). The antisense molecules may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. Examples, which may be used, include engineered hammerhead motif ribozyme molecules that can be specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding the proteins of the invention and homologous proteins. Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target gene containing the cleavage site may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Nucleic acid effector molecules, e.g. antisense molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences. Such DNA sequences may be incorporated into a variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize antisense RNA constitutively or inducibly can be introduced into cell lines, cells or tissues. RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule or modifications in the nucleobase, sugar and/or phosphate moieties, e.g. the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of non-traditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection and by liposome injections may be achieved using methods, which are well known in the art. Any of the therapeutic methods described above may be applied to any suitable subject including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

An additional embodiment of the invention relates to the administration of a pharmaceutical composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of the nucleic acids and the proteins of the invention and homologous nucleic acids or proteins, antibodies to the proteins of the invention and homologous proteins, mimetics, agonists, antagonists or inhibitors of the proteins of the invention and homologous proteins or nucleic acids. The compositions may be administered alone or in combination with at least one other agent, such as stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone or in combination with other agents, drugs or hormones. The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries, which facilitate processing of the active compounds into preparations, which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing Co., Easton, Pa.).

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art. For any compounds, the therapeutically effective does can be estimated initially

either in cell culture assays, e.g., of preadipocyte cell lines or in animal models, usually mice, rabbits, dogs or pigs. The animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans. A therapeutically effective dose refers to that amount of active ingredient, for example the nucleic acids or the proteins of the invention or fragments thereof, or antibodies of the proteins of the invention and homologous proteins, which is sufficient for treating a specific condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., ED50 (the dose therapeutically effective in 50% of the population) and LD50 (the dose lethal to 50% of the population). The dose ratio between therapeutic and toxic effects is the therapeutic index, and it can be expressed as the ratio, LD50/ED50. Pharmaceutical compositions, which exhibit large therapeutic indices, are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage varies within this range depending upon the dosage from employed, sensitivity of the patient, and the route of administration. The exact dosage will be determined by the practitioner, in light of factors related to the subject that requires treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors, which may be taken into account, include the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week or once every two weeks depending on half-life and clearance rate of the particular formulation. Normal dosage amounts may vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the

route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

In another embodiment, antibodies which specifically bind to the proteins may be used for the diagnosis of conditions or diseases characterized by or associated with over- or underexpression of the proteins of the invention and homologous proteins or in assays to monitor patients being treated with the proteins of the invention and homologous proteins, or effectors thereof, e.g. agonists, antagonists, or inhibitors. Diagnostic assays include methods which utilize the antibody and a label to detect the protein in human body fluids or extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by joining them, either covalently or non-covalently, with a reporter molecule. A wide variety of reporter molecules which are known in the art may be used several of which are described above.

A variety of protocols including ELISA, RIA, and FACS for measuring proteins are known in the art and provide a basis for diagnosing altered or abnormal levels of gene expression. Normal or standard values for gene expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, preferably human, with antibodies to the protein under conditions suitable for complex formation. The amount of standard complex formation may be quantified by various methods, but preferably by photometric means. Quantities of protein expressed in control and disease, samples e.g. from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides specific for the proteins of the invention and homologous proteins may be used for diagnostic purposes. The polynucleotides, which may be used, include oligonucleotide sequences, antisense RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantitate gene expression in biopsied tissues in which gene expression may be correlated with disease. The diagnostic assay may be used to distinguish between absence, presence, and excess gene expression, and to monitor regulation of protein levels during therapeutic intervention.

In one aspect, hybridization with probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding the proteins of the invention and homologous proteins or closely related molecules, may be used to identify nucleic acid sequences which encode the respective protein. The hybridization probes of the subject invention may be DNA or RNA and are preferably derived from the nucleotide sequence of the polynucleotide encoding the proteins of the invention or from a genomic sequence including promoter, enhancer elements, and introns of the naturally occurring gene. Hybridization probes may be labeled by a variety of reporter groups, for example, radionuclides such as ^{32}P or ^{35}S or enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences specific for the proteins of the invention and homologous nucleic acids may be used for the diagnosis of conditions or diseases, which are associated with the expression of the proteins. Examples of such conditions or diseases include, but are not limited to, pancreatic diseases and disorders, including diabetes. Polynucleotide sequences specific for the proteins of the invention and homologous proteins may also be used to monitor the progress of patients receiving treatment for pancreatic diseases and disorders, including diabetes. The polynucleotide sequences may be used qualitative or quantitative assays,

e.g. in Southern or Northern analysis, dot blot or other membrane-based technologies; in PCR technologies; or in dip stick, pin, ELISA or chip assays utilizing fluids or tissues from patient biopsies to detect altered gene expression.

In a particular aspect, the nucleotide sequences specific for the proteins of the invention and homologous nucleic acids may be useful in assays that detect activation or induction of various metabolic diseases or dysfunctions, for example, obesity, diabetes mellitus, eating disorder, cachexia, hypertension, coronary heart disease, hypercholesterolemia (dyslipidemia), and gallstones. The nucleotide sequences may be labeled by standard methods, and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantitated and compared with a standard value. If the amount of signal in the biopsied or extracted sample is significantly altered from that of a comparable have hybridized with nucleotide sequences in the sample, and the presence of altered levels of nucleotide sequences encoding the proteins of the invention and homologous proteins in the sample indicates the presence of the associated disease. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials or in monitoring the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disease associated with expression of the proteins of the invention and homologous proteins, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence or a fragment thereof, which is specific for the nucleic acids encoding the proteins of the invention and homologous nucleic acids, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by

comparing the values obtained from normal subjects with those from an experiment where a known amount of a substantially purified polynucleotide is used. Standard values obtained from normal samples may be compared with values obtained from samples from patients who are symptomatic for disease. Deviation between standard and subject values is used to establish the presence of disease. Once disease is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to evaluate whether the level of expression in the patient begins to approximate that, which is observed in the normal patient. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to metabolic diseases such as described above the presence of an unusual amount of transcript in biopsied tissue from an individual may indicate a predisposition for the development of the disease or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the metabolic diseases and disorders.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding the proteins of the invention and homologous proteins may involve the use of PCR. Such oligomers may be chemically synthesized, generated enzymatically or produced from a recombinant source. Oligomers will preferably consist of two nucleotide sequences, one with sense orientation (5prime.fwdarw.3prime) and another with antisense (3prime.rarw.5prime), employed under optimized conditions for identification of a specific gene or condition. The same two oligomers, nested sets of oligomers or even a degenerate pool of oligomers may be employed under less stringent conditions for detection and/or quantification of closely related DNA or RNA sequences.

In another embodiment of the invention, the nucleic acid sequences may also be used to generate hybridization probes, which are useful for mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome or to a specific region of the chromosome using well known techniques. Such techniques include FISH, FACS or artificial chromosome constructions, such as yeast artificial chromosomes, bacterial artificial chromosomes, bacterial P1 constructions or single chromosome cDNA libraries as reviewed in Price, C. M. (1993) Blood Rev. 7:127-134, and Trask, B. J. (1991) Trends Genet. 7:149-154. FISH (as described in Verma et al. (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York, N.Y.). The results may be correlated with other physical chromosome mapping techniques and genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of the gene encoding the proteins of the invention on a physical chromosomal map and a specific disease or predisposition to a specific disease, may help to delimit the region of DNA associated with that genetic disease.

The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals. An analysis of polymorphisms, e.g. single nucleotide polymorphisms may be carried out. Further, in situ hybridization of chromosomal preparations and physical mapping techniques such as linkage analysis using established chromosomal markers may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms or parts thereof, by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular

genomic region, for example, AT to 11q22-23 (Gatti, R. A. et al. (1988) Nature 336:577-580), any sequences mapping to that area may represent associated or regulatory genes for further investigation. The nucleotide sequences of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc. among normal, carrier or affected individuals.

In another embodiment of the invention, the proteins of the invention, their catalytic or immunogenic fragments or oligopeptides thereof, an in vitro model, a genetically altered cell or animal, can be used for screening libraries of compounds in any of a variety of drug screening techniques. One can identify effectors, e.g. receptors, enzymes, proteins, ligands, or substrates that bind to, modulate or mimic the action of one or more of the proteins of the invention. The protein or fragment thereof employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellulary. The formation of binding complexes, between the proteins of the invention and the agent tested, may be measured. Agents can also be identified, which, either directly or indirectly, influence the activity of the protein of the invention. For example the phosphatase activity of the proteins of the invention could be measured in vitro by using recombinantly expressed and purified Prl-1 or fragments thereof by making use of artificial substrates well known in the art, i.e. but not exclusively DiFMUP (Molecular Probes, Eugene, Oregon), which are converted to fluorophores or chromophores upon dephosphorylation. Alternatively, the dephosphorylation of physiological substrates of the phosphatases could be measured by making use of any of the well known screening technologies suitable for the detection of the phosphorylation status of their physiological substrates. For example, but not exclusively, the phosphorylation status of peptides derived from their physiological substrates can be monitored by binding of phospho-side specific antibodies resulting in an increase of the polarization of the complex.

In addition activity of Prl-1 against its physiological substrate(s) or derivatives thereof could be measured in cell-based assays. Agents may also interfere with posttranslational modifications of the protein, such as phosphorylation and dephosphorylation, farnesylation, palmitoylation, acetylation, alkylation, ubiquitination, proteolytic processing, subcellular localization and degradation. Moreover, agents could influence the dimerization or oligomerization of the proteins of the invention or, in a heterologous manner, of the proteins of the invention with other proteins, for example, but not exclusively, docking proteins, enzymes, receptors, or translation factors. Agents could also act on the physical interaction of the proteins of this invention with other proteins, which are required for protein function, for example, but not exclusively, their downstream signaling.

Methods for determining protein-protein Interaction are well known in the art. For example binding of a fluorescently labeled peptide derived from the interacting protein to the protein of the Invention, or vice versa, could be detected by a change in polarisation. In case that both binding partners, which can be either the full length proteins as well as one binding partner as the full length protein and the other just represented as a peptide are fluorescently labeled, binding could be detected by fluorescence energy transfer (FRET) from one fluorophore to the other. In addition, a variety of commercially available assay principles suitable for detection of protein-protein Interaction are well known In the art, for example but not exclusively AlphaScreen (PerkinElmer) or Scintillation Proximity Assays (SPA) by Amersham. Alternatively, the interaction of the proteins of the invention with cellular proteins could be the basis for a cell-based screening assay, in which both proteins are fluorescently labeled and interaction of both proteins is detected by analysing cotranslocation of both proteins with a cellular imaging reader, as has been developed for example, but not exclusively, by Cellomics or EvotecOAI. In all cases the two or more binding partners can be different proteins with one being the protein of the invention, or in case of dimerization and/or oligomerization the protein of

the invention itself. Proteins of the invention, for which one target mechanism of interest, but not the only one, would be such protein/protein Interactions are Prl-1.

Of particular interest are screening assays for agents that have a low toxicity for mammalian cells. The term "agent" as used herein describes any molecule, e.g. protein or pharmaceutical, with the capability of altering or mimicking the physiological function of one or more of the proteins of the invention. Candidate agents encompass numerous chemical classes, though typically they are organic molecules, preferably small organic compounds having a molecular weight of more than 50 and less than about 2,500 Daltons. Candidate agents comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups. The candidate agents often comprise carbocyclic or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups.

Candidate agents are also found among biomolecules including peptides, saccharides, fatty acids, steroids, purines, pyrimidines, nucleic acids and derivatives, structural analogs or combinations thereof. Candidate agents are obtained from a wide variety of sources including libraries of synthetic or natural compounds. For example, numerous means are available for random and directed synthesis of a wide variety of organic compounds and biomolecules, including expression of randomized oligonucleotides and oligopeptides. Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available or readily produced. Additionally, natural or synthetically produced libraries and compounds are readily modified through conventional chemical, physical and biochemical means, and may be used to produce combinatorial libraries. Known pharmacological agents may be subjected to directed or

random chemical modifications, such as acylation, alkylation, esterification, amidification, etc. to produce structural analogs. Where the screening assay is a binding assay, one or more of the molecules may be joined to a label, where the label can directly or indirectly provide a detectable signal.

Another technique for drug screening, which may be used, provides for high throughput screening of compounds having suitable binding affinity to the protein of interest as described in published PCT application WO84/03564. In this method, as applied to the proteins of the invention large numbers of different small test compounds, e.g. aptamers, peptides, low-molecular weight compounds etc., are provided or synthesized on a solid substrate, such as plastic pins or some other surface. The test compounds are reacted with the proteins or fragments thereof, and washed. Bound proteins are then detected by methods well known in the art. Purified proteins can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support. In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding the protein specifically compete with a test compound for binding the protein. In this manner, the antibodies can be used to detect the presence of any peptide, which shares one or more antigenic determinants with the protein.

Finally, the invention also relates to a kit comprising at least one of

(a) an Prl-1 nucleic acid molecule or a fragment thereof;

(b) a vector comprising the nucleic acid of (a);

(c) a host cell comprising the nucleic acid of (a) or the vector of (b);

(d) a polypeptide encoded by the nucleic acid of (a);

(e) a fusion polypeptide encoded by the nucleic acid of (a);

(f) an antibody, an aptamer or another receptor against the nucleic acid of (a) or the polypeptide of (d) or (e) and

(g) an anti-sense oligonucleotide of the nucleic acid of (a).

The kit may be used for diagnostic or therapeutic purposes or for screening applications as described above. The kit may further contain user instructions.

The Figures show:

FIGURE 1 shows the content of energy storage metabolites (ESM; triglyceride (TG) and glycogen) of Drosophila Prl-1 (GadFly Accession Number CG4993) mutants. Shown is the change of triglyceride content of HD-EP(2)20261 flies caused by integration of the P-vector into the into the annotated transcription unit (column 3) in comparison to controls containing more than 2000 fly lines of the proprietary EP collection ('HD-control (TG)', column 1) and wildtype controls determined in more than 80 independent assays (referred to as 'WT-control (TG)' column 2). Also shown is the change of glycogen content of HD-EP(2)20261 flies caused by integration of the P-vector the into the annotated transcription unit (column 5) in comparison to controls (referred to as 'control (glycogen)' column 4).

FIGURE 2 shows the molecular organization of the mutated Prl-1 (Gadfly Accession Number CG4993) gene locus.

FIGURE 3 shows the nucleic acid and amino acid sequences of the proteins of the invention.

Figure 3A shows the nucleic acid sequence of human Prl-1 (SEQ ID NO: 1). Figure 3B shows the amino acid sequence (one-letter code) of human Prl-1 (SEQ ID NO: 2).

Figure 3C shows the nucleic acid sequence of human Prl-2, transcript variant 1 (SEQ ID NO: 3)

Figure 3D shows the amino acid sequence (one-letter code) of human Prl-2, transcript variant 1 (SEQ ID NO: 4).

Figure 3E shows the nucleic acid sequence of human Prl-2, transcript variant 2 (SEQ ID NO: 5)

Figure 3F shows the amino acid sequence (one-letter code) of human Prl-2, transcript variant 2 (SEQ ID NO: 6).

Figure 3G shows the nucleic acid sequence of human Prl-2, transcript variant 3 (SEQ ID NO: 7)

Figure 3H shows the amino acid sequence (one-letter code) of human Prl-2, transcript variant 3 (SEQ ID NO: 8).

Figure 3I shows the nucleic acid sequence of human Prl-3 (SEQ ID NO: 9)

Figure 3J shows the amino acid sequence (one-letter code) of human Prl-3 (SEQ ID NO: 10).

Figure 4 shows the expression of Prl-1 in different mammalian models

Figure 4A. Real-time PCR analysis of Prl-1 expression in wildtype mouse tissues. The relative RNA-expression is shown on the Y-axis, the tissues tested are given on the X-axis. (WAT = white adipose tissue, light grey columns; BAT = brown adipose tissue, dark grey columns)

Figure 4B shows real-time PCR analysis of Prl-1 expression in metabolically active tissues of wildtype mice, of mice on high-fat-diet (HFD), of fasted mice (Fasted), of genetically obese mice (ob/ob mice and db/db mice). The relative RNA-expression is shown on the Y-axis, the mouse models tested are given on the X-axis. (WAT = white adipose tissue, light grey columns; BAT = brown adipose tissue, dark grey columns)

Figure 5 shows in vitro assays for the determination of triglyceride and glycogen levels in cells overexpressing Prl-1.

Figure 5A shows an increase in triglyceride levels in cells overexpressing Prl-1. The Y-axis shows cellular triglyceride (shown as microg triglyceride per mg protein) levels and the X-axis shows days of cell differentiation. Measurements from cells overexpressing Prl-1 are shown as dark grey columns, control cells (empty vector) are shown as light grey columns. Triglyceride levels and controls are shown for two different sets of samples.

Figure 5B shows an increase in glycogen levels in cells overexpressing Prl-1. The Y-axis shows glycogen levels (shown as microg glycogen per mg protein) and the X-axis shows days of cell differentiation. Measurements from cells overexpressing Prl-1 are shown as dark grey columns, control cells (empty vector) are shown as light grey columns. Glycogen levels and controls are shown for two different sets of samples.

The examples illustrate the invention:

Example 1: Measurement of energy storage metabolites (ESM) contents in Drosophila

Mutant flies are obtained from a fly mutation stock collection. The flies are grown under standard conditions known to those skilled in the art. In the course of the experiment, additional feedings with bakers yeast (Saccharomyces cerevisiae) are provided for the EP-line HD-EP20261. The average change of triglyceride and glycogen (herein referred to as energy storage metabolites, ESM) content of Drosophila containing the EP-vector as homozygous viable integration was investigated in comparison to control flies, respectively (see FIGURE 1). For determination of ESM content, flies were incubated for 5 min at 90°C in an aqueous buffer using a waterbath, followed by hot extraction. After another 5 min incubation at 90°C and mild centrifugation, the triglyceride content of the flies extract was determined using Sigma Triglyceride (INT 336-10 or -20) assay by measuring changes in the optical density according to the manufacturer's protocol, and the glycogen content of the flies extract was determined using Roche (Starch UV-method Cat. No. 0207748) assay by measuring changes in the optical density according to the manufacturer's protocol. As a reference the protein content of the same extract was measured using BIO-RAD DC Protein Assay according to the manufacturer's protocol. These experiments and assays were repeated several times.

The average triglyceride level (µg triglyceride/µg protein) of 2108 fly lines of the proprietary EP-collection (referred to as 'HD-control (TG)') is shown as 100% in the first column in FIGURE 1. The average triglyceride level (µg triglyceride/µg protein) of drosophila wildtype strain Oregon R flies determined in 84 independent assays (referred to as 'WT-control (TG)') is shown as 102% in the second column in FIGURE 1. The average glycogen level (µg glycogen/µg protein) of an internal assay control consisting of two different wildtype strains and an inconspicuous EP-line of the HD stock collection (referred to as 'control (glycogen)') is shown as 100% in the fourth column in FIGURE 1. Standard deviations of the measurements are shown as thin bars. HD-20261 homozygous flies show constantly a higher triglyceride content (µg triglyceride/µg protein) than the controls (column 3 in FIGURE 1, 'HD-20261 (TG)'). HD-20261 homozygous flies also show a lower glycogen content (µg triglyceride/µg protein) than the controls (column 5 in FIGURE 1, 'HD-20261 (glycogen)'). Therefore, the loss of gene activity is responsible for changes in the metabolism of the energy storage triglycerides.

Example 2: Identification of a Drosophila gene responsible for changes in triglyceride and glycogen levels

Genomic DNA sequences were isolated that are localized directly adjacent to the EP vector (herein HD-EP(2)20261) integration. Using those isolated genomic sequences public databases like Berkeley Drosophila Genome Project (GadFly) were screened thereby confirming the homozygous viable integration site of the HD-EP(2)20261 vector at base pair 9 of the transcript variant CG4993-RB of the gene Prl-1 in sense orientation. FIGURE 2 shows the molecular organization of this gene locus. The chromosomal localization site of integration of the vector of HD-EP(2)20261 is at gene locus 2L, 35E2 (according to Flybase) or 2L, 35F1 (according toGadfly release 3). In FIGURE 2, genomic DNA sequence is represented by the assembly as a black arrow in middle of the figure that

includes the integration site of HD-EP(2)20261. Ticks represent the length in basepairs of the genomic DNA (1000 base pairs per tick). Dark grey bars in the lower half of the figure, linked by dark grey lines represent cDNAs of the predicted genes (as predicted by the Berkeley Drosophila Genome Project, GadFly release 3). Predicted exons of the Drosophila cDNA of the gene Prl-1 (GadFly Accession Number CG4993) are shown as dark grey bars and predicted introns as slim grey lines in the lower half of the figure and are labeled. The integation site of HD-EP20261 is indicated with a black triangle within the first exon of the Prl-1 perdicted cDNA. The corresponding expressed Sequence tags (ESTs) are shown as light grey bars below the two cDNA transcript variants. Therefore, expression of the cDNA encoding Prl-1 could be affected by integration of the vector of line HD-EP(2)20261, leading to a change in the amount of energy storage triglycerides.

Table 1 is summarizing the data of our molecular analysis of the Drosophila protein identified in this invention as being involved in the regulation of the metabolism.

Table 1. Molecular analysis of Drosophila Prl-1

Analysis	Prl-1
Genetic interaction:	not described (Flybase)
Protein	prenylated protein tyrosine phosphatase (Flybase)
Protein domains	Dual specificity protein phosphatase, Tyrosine specific protein phosphatase and dual specificity protein phosphatase family, (Phosphotyrosine protein) phosphatases II (Flybase)
InterPro analysis	Tyrosine specific protein phosphatase and dual specificity protein phosphatase (IPR000387)
Drosophila functional data	not described (Flybase)
Locus	2L, 35E2 (Flybase); 2L, 35F1 (Gadfly release 3)
Ests	many including RE55984 (Gadfly release 3)
CDNA	AA140816 (625 bp mRNA, 1998), AA392622 (577 bp mRNA, 2001), AA538688 (511 bp mRNA, 2001), AA540959 (443 bp mRNA, 2001), AA696442 (392 bp mRNA, 2001), AA803899 (530 bp mRNA, 2001), AA821150 (543 bp mRNA, 2001), AA821153 (552 bp mRNA, 2001), AA979376 (698 bp mRNA, 2001), AA979397 (785 bp mRNA, 2001), AA979413 (708 bp mRNA, 2001), AA990900 (625 bp mRNA, 2001), AF063902 (537 bp mRNA, 1998; protein:AAC16552), AF390535 (531 bp mRNA, 2001; protein:AAL26988), AI063355 (713 bp mRNA, 2001), AY071505 (3167 bp mRNA, 2001; protein:AAL49127), BI369716 (607 bp mRNA, 2001) (Flybase)
genomic DNA	AE003415 (299935 bp DNA, 2000; protein:AAF44989), AE003650 (264581 bp DNA, 2000; protein:AAF53506) (Flybase)
NCBI locus ID	34952, Dm Prl-1, 35E2. Aliases: PRL, CG4993, CT16026, BG:DS07473.3 RefSeq: NM_135936 Nucleotide: AE003415, AE003650, AQ025903, AQ025971, AQ073277, AQ074056, AA140816, AA392622, AA538688, AA540959, AA696442, AA803899, AA821150, AA821153, AA979376, AA979397, AA979413, AA990900, AF063902, AF390535, AY071505 Protein: NP_609780, AAF44989, AAF53506, AAC16552, AAL26988, AAL49127
Drosophila mutations & mutants	There are 5 recorded alleles: 4 classical mutants (1 available from the public stock centers) and 1 wild-type. (Flybase)

Example 3: Identification of the human Prl-1 homologous proteins

Prl-1 homologous proteins and nucleic acid molecules coding therefore are obtainable from insect or vertebrate species, e.g. mammals or birds. Particularly preferred are nucleic acids comprising Drosophila Prl-1 or human Prl-1 and homologs. Sequences homologous to Drosophila Prl-1 were identified using the publicly available program BLASTP 2.2.3 of the non-redundant protein data base of the National Center for Biotechnology Information (NCBI) (see, Altschul et al., 1997, Nucleic Acids Res. 25:3389-3402). Table 2 shows the best human homologs of the Drosophila Prl-1 gene.

Table 2. Human homologous proteins to Drosophila Prl-1 protein

NCBI (National Center for Biotechnology Information) human locus identification (ID): 7803, Hs PTP4A1, protein tyrosine phosphatase type IVA, member 1, 6q12

* Aliases: HH72, PRL1, Prl-1, PTP4A2, PTPCAAX1, PTP(CAAX1)
* OMIM: 601585
* RefSeq[R]: NM_003463
* Nucleotide: AF051160, AJ420505, BC023975, BI222469, U48296, U69701
* Protein: NP_003454, AAC39836, AAH23975, AAB40597, AAB09080

NCBI (National Center for Biotechnology Information) human locus identification (ID): 8073, Hs PTP4A2, protein tyrosine phosphatase type IVA, member 2, 1p35

* Aliases: HH13, OV-1, PRL2, HH7-2, PRL-2, PTP4A, HU-PP-1, PTPCAAX2, ptp-IV1a, ptp-IV1b
* OMIM: 601584
* RefSeq[R]: NM_003479, NM_080391, NM_080392
* Nucleotide: L48722, L48723, L48937, AF208850, BI552091, L39000, U14603, U48297
* Protein: NP_003470, NP_536316, NP_536317, AAB42169, AAB42170, AAB39331, AAF64264, AAB59575, AAA90979, AAB40598

NCBI (National Center for Biotechnology Information) human locus identification (ID): 11156, Hs PTP4A3, protein tyrosine phosphatase type IVA, member 3

* Aliases: PRL3, PRL-3, PRL-R
* OMIM: 606449
* RefSeq[R]: NM_007079, NM_032611
* Nucleotide: AF041434, AI742376, BC003105, BE778111, U87168
* Protein: NP_009010, NP_116000, AAC29314, AAH03105, AAB47560

* * * * * * * * * * * * * * * * * *

The mouse homologous cDNAs encoding the polypeptides of the invention were identified as GenBank Accession Numbers NM_011200, XM_123656, XM_135289, NM_008974, NM_008975.

Example 4: Prl-1 mRNA expression in mammalian (mouse) tissues (Figure 4)

To analyse the expression of the Prl-1 mRNA disclosed in this invention in mammalian tissues, several mouse strains (preferably mice strains C57Bl/6J, C57Bl/6 ob/ob and C57Bl/KS db/db which are standard model systems in obesity and diabetes research) were purchased from Harlan Winkelmann (33178 Borchen, Germany) and maintained under constant temperature (preferably 22°C), 40 per cent humidity and a light / dark cycle of preferably 14 / 10 hours. The mice were fed a standard chow (for example, from ssniff Spezialitäten GmbH, order number ssniff M-Z V1126-000). For the fasting experiment ("fasted wild type mice"), wild type mice were starved for 48 h without food, but only water supplied ad libitum. (see, for example, Schnetzler et al. J Clin Invest 1993 Jul;92(1):272-80, Mizuno et al. Proc Natl Acad Sci U S A 1996 Apr 16;93(8):3434-8). Animals were sacrificed at an age of 6 to 8 weeks. The animal tissues were isolated according to standard procedures known to those skilled in the art, snap frozen in liquid nitrogen and stored at -80°C until needed.

For analyzing the role of the proteins disclosed in this invention in the in vitro differentiation of different mammalian cell culture cells for the conversion of pre-adipocytes to adipocytes, mammalian fibroblast (3T3-L1) cells (e.g., Green & Kehinde, Cell 1: 113-116, 1974) were obtained from the American Tissue Culture Collection (ATCC, Hanassas, VA, USA; ATCC- CL 173). Alternatively, a mammalian fibroblast TA1 cell line, a murine preadipocyte line derived from T101/2 mouse embryo fibroblasts (Chapman et al., 1984, J Biol Chem 259(24):15548-55), was used. 3T3-L1 or TA-1cells were maintained as fibroblasts and differentiated into adipocytes as described in the prior art (e.g., Qiu. et al., J. Biol. Chem. 276:11988-95, 2001; Slieker et al., BBRC 251: 225-9, 1998). In brief, cells were plated in DMEM/10% FCS (Invitrogen, Karlsruhe, Germany) at 50,000 cells/well in duplicates in 6-well plastic dishes and cultured in a humidified atmosphere of 5% CO_2 at 37°C. At confluence (defined as day 0: d0) cells were transferred to serum-free (SF) medium, containing DMEM/HamF12 (3:1; Invitrogen), Fetuin (300 µg/ml; Sigma, Munich, Germany), Transferrin (2 µg/ml; Sigma), Pantothenate (17 µM; Sigma), Biotin (1 µM; Sigma), and EGF (0.8 nM; Hoffmann-La Roche, Basel, Switzerland). Differentiation was induced by adding Dexamethasone (DEX; 1 µM; Sigma), 3-Methyl-Isobutyl-1-Methylxanthine (MIX; 0.5 mM; Sigma), and bovine Insulin (5 µg/ml; Invitrogen). Four days after confluence (d4), cells were kept in SF medium, containing bovine Insulin (5 µg/ml) until differentiation was completed. At various time points of the differentiation procedure, beginning with day 0 (day of confluence) and day 2 (hormone addition; for example, dexamethasone and 3-isobutyl-1-methylxanthine), up to 10 days of differentiation, suitable aliquots of cells were taken every two days.

TaqMan Analysis of the Prl-1 mRNA expression (Figure 4A and Figure 4B) RNA was isolated from mouse tissues or cell culture cells using Trizol Reagent (for example, from Invitrogen, Karlsruhe, Germany) and further purified with the RNeasy Kit (for example, from Qiagen, Germany) in

combination with an DNase-treatment according to the instructions of the manufacturers and as known to those skilled in the art. Total RNA was reverse transcribed (preferably using Superscript II RNaseH- Reverse Transcriptase, from Invitrogen, Karlsruhe, Germany) and subjected to Taqman analysis preferably using the Taqman 2xPCR Master Mix (from Applied Biosystems, Weiterstadt, Germany; the Mix contains according to the Manufacturer for example AmpliTaq Gold DNA Polymerase, AmpErase UNG, dNTPs with dUTP, passive reference Rox and optimized buffer components) on a GeneAmp 5700 Sequence Detection System (from Applied Biosystems, Weiterstadt, Germany).

The following prime/probe pairs were used for the TaqMan analysis (GenBank Accession Number U84411 for the mouse Prl-1 sequence):

Mouse Prl-1 reverse primer (Seq ID NO:11)
5'GCTGTATTGCTGTCCATTGTGTC-3';
Mouse Prl-1 reverse primer (Seq ID NO:12)
5'TCCACCTTCAATTAATGCTAGGG-3';
Mouse Prl-1 Taqman probe (Seq ID NO:13)
(5/6-FAM) CAGGCCTTGGCAGAGCTCCGG (5/6-TAMRA)

Expression profiling studies confirm the particular relevance of Prl-1 as regulator of energy metabolism in mammals. Taqman analysis revealed that Prl-1 is expressed in varieties of tissues. However, it is highly expressed in metabolic active tissue such as brown adipose tissue (BAT) compared to other tissue types in wild type mouse as depicted in Figure 4A (pancreas was used as a reference tissue). This high expression of Prl-1 in BAT confirms an essential role of Prl-1 in adipogenesis.

Further, we show that the mammalian homologue of the Drosophilia Prl-1 gene is regulated by fasting and HFD and by genetically induced obesity.

In this invention, we used mouse models of insulin resistance and/or diabetes, such as mice carrying gene knockouts in the leptin pathway (for example, ob (leptin) or db (leptin receptor/ligand) mice) to study the expression of the protein of the invention. Such mice develop typical symptoms of diabetes, show hepatic lipid accumulation and frequently have increased plasma lipid levels (see Bruning et al, 1998, Mol. Cell. 2:449-569). We found, for example, that the expression of Prl-1 is strongly upregulated in metabolic active tissue BAT compared to white adipose tissue (WAT) in genetically induced obese mice (ob-ob and db-db) (Figure 4B). RNA expression was also examined in susceptible wild type mice (for example C57Bl/6) that show symptoms of diabetes, lipid accumulation, and high plasma lipid levels, if fed a high fat diet (HFD). In those mice, the expression of Prl-1 is significantly upregulated in brown adipose tissue supporting a hypothesis that Prl-1 is a modulator of adipogenesis. In addition, Prl-1 is upregulated in fasted mice (Figure 4B).

Example 5: Assays for the determination of triglyceride and glycogen storage (Figure 5)

Retroviral infection of preadipocytes

Packaging cells were transfected with retroviral plasmids pLPCX carrying mouse Prl-1 transgene and a selection marker using calcium phosphate procedure. Control cells were infected with pLPCX carrying no transgene. Briefly, exponentially growing packaging cells were seeded at a density of 350,000 cells per 6-well in 2 ml DMEM + 10 % FCS one day before transfection. 10 min before transfection chloroquine was added directly to the overlying medium (25 μM end concentration). A 250 μl transfection mix consisting of 5 μg plasmid-DNA (candidate:helper-virus in a 1:1 ratio) and 250 mM $CaCl_2$ was prepared in a 15 ml plastic tube. The same volume of 2 x HBS (280 μM NaCl, 50 μM HEPES, 1.5 mM Na_2HPO_4, pH 7.06) was added and air bubbles were injected into the mixture for 15 sec. The

transfection mix was added drop wise to the packaging cells, distributed and the cells were incubated at 37°C, 5 % CO_2 for 6 hours. The cells were washed with PBS and the medium was exchanged with 2 ml DMEM + 10 % CS per 6-well. One day after transfection the cells were washed again and incubated for 2 days of virus collection in 1 ml DMEM + 10 % CS per 6-well at 32°C, 5 % CO_2. The supernatant was then filtered through a 0.45 μm cellulose acetate filter and polybrene (end concentration 8 μg/ml) was added. Mammalian fibroblast (3T3-L1) cells in a sub-confluent state were overlaid with the prepared virus containing medium. The infected cells were selected for 1 week with 2 μg/ml puromycin. Following selection the cells were checked for transgene expression by western blot and immunofluorescence. Over expressing cells were seeded for differentiation.

3T3-L1 cells were maintained as fibroblasts and differentiated into adipocytes as described in the prior art and supra. For analysing the role of the proteins disclosed in this invention in the in vitro assays for the determination of triglyceride storage, synthesis and transport were performed.

Preparation of cell lysates for analysis of metabolites

Starting at confluence (D0), cell media was changed every 48 hours. Cells and media were harvested 8 hours prior to media change as follows. Media was collected, and cells were washed twice in PBS prior to lyses in 600 μl HB-buffer (0.5% polyoxyethylene 10 tridecylethane, 1 mM EDTA, 0.01M NaH_2PO_4, pH 7.4). After inactivation at 70°C for 5 minutes, cell lysates were prepared on Bio 101 systems lysing matrix B (0.1 mm silica beads; Q-Biogene, Carlsbad, USA) by agitation for 2 x 45 seconds at a speed of 4.5 (Fastprep FP120, Bio 101 Thermosavant, Holbrock, USA). Supernatants of lysed cells were collected after centrifugation at 3000 rpm for 2 minutes, and stored in aliquots for later analysis at -80°C.

Changes in cellular triglyceride levels during adipogenesis (Figure 5A)

Cell lysates and media were simultaneously analysed in 96-well plates for total protein and triglyceride content using the Bio-Rad DC Protein assay reagent (Bio-Rad, Munich, Germany) according to the manufacturer's instructions and a modified enzymatic triglyceride kit (GPO-Trinder; Sigma) briefly final volumes of reagents were adjusted to the 96-well format as follows: 10 µl sample was incubated with 200 µl reagent A for 5 minutes at 37°C. After determination of glycerol (initial absorbance at 540 nm), 50 µl reagent B was added followed by another incubation for 5 minutes at 37°C (final absorbance at 540 nm). Glycerol and triglyceride concentrations were calculated using a glycerol standard set (Sigma) for the standard curve included in each assay.

As shown in Figure 5A, we found that in Prl 1 overexpressing cells cellular triglyceride levels were increased on day 12 of differentiation when compared to control cells which were transduced with empty vector. An increase in triglyceride levels of about 20% in these experiments is significant. When we overexpress known regulators of adipogenesis such as for example PPAR gamma-1 in 3T3-L1 cells, we repeatedly observe increases in triglyceride content between 20 and 30% as compared to control cells.

Changes in cellular glycogen levels during adipogenesis (Figure 5B)

Cell lysates and media were simultaneously analysed in triplicates in 96-well plates for total protein and glycogen content using the Bio-Rad DC Protein assay reagent (Bio-Rad, Munich, Germany) according to the manufacturer's instructions and an enzymatic starch kit from Hoffmann-La Roche (Basel, Switzerland). 10 µl samples were incubated with 20 µl amyloglucosidase solution for 15 minutes at 60°C to digest glycogen to glucose. The glucose is further metabolised with 100 µl distilled water and 100 µl of enzyme cofactor buffer and 12 µl of enzyme buffer (hexokinase and glucose phosphate dehydrogenase). Background glucose levels are

determined by subtracting values from a duplicate plate without the amyloglucosidase. Final absorbance is determined at 340 nm. HB-buffer as blank, and a standard curve of glycogen (Hoffmann-La Roche) were included in each assay. Glycogen content in samples were calculated using a standard curve.

As shown in Figure 5B, we found that in Prl-1 overexpressing cells cellular glycogen levels were increased throughout adipogenesis. Glycogen levels in cells are more variable than triglyceride levels because the turnover of glycogen is higher. Glucose is taken up by the cells rapidly and stored in the form of glycogen. This energy storage is then used as a first quick response to the metabolic demands of the cell. On day 12 of differentiation intracellular glycogen levels are increased by more than 100% indicating that the Prl-1 effects central metabolic pathways in the cell.

Example 6: Generation and analysis of Prl-1 transgenic mice

Generation of the transgenic animals

Mouse Prl-1 cDNA was isolated from mouse brown adipose tissue (BAT) using standard protocols as known to those skilled in the art. The cDNA was amplified by RT-PCR and point mutations were introduced into the cDNA.

The resulting mutated cDNA was cloned into a suitable transgenic expression vector. The transgene was microinjected into the male pronucleus of fertilized mouse embryos (preferably strain C57/BL6/CBA F1 (Harlan Winkelmann). Injected embryos were transferred into pseudo-pregnant foster mice. Transgenic founders were detected by PCR analysis. Two independent transgenic mouse lines containing the construct were established and kept on a C57/BL6 background. Briefly, founder animals were backcrossed with C57/BL6 mice to generate F1 mice for analysis. Transgenic mice were continously bred onto the C57/Bl6

background. The expression of the protein of the invention can be analyzed by taqman analysis as described above, and further analysis of the mice can be done as known to those skilled in the art.

0 3. Dez. 2002

EPO - Munich
69
0 3. Dez. 2002

Claims

1. A pharmaceutical composition comprising a nucleic acid molecule of the protein tyrosine phosphatase family or a polypeptide encoded thereby or a fragment or a variant of said nucleic acid molecule or said polypeptide or an effector of said nucleic acid molecule or polypeptide together with pharmaceutically acceptable carriers, diluents and/or adjuvants.

2. The composition of claim 1, wherein the nucleic acid molecule is a vertebrate or insect Prl-1 nucleic acid, particulary encoding the human Prl-1 homologs (such as human Prl-1, Prl-2, or Prl-3 protein), and/or a nucleic molecule which is complementary thereto or a fragment thereof or a variant thereof.

3. The composition of claim 1 or 2, wherein said nucleic acid molecule
   (a) hybridizes at 50°C in a solution containing 1 x SSC and 0.1% SDS to a nucleic acid molecule as defined in claim 2 and/or a nucleic acid molecule which is complementary thereto;
   (b) it is degenerate with respect to the nucleic acid molecule of (a),
   (c) encodes a polypeptide which is at least 85%, preferably at least 90%, more preferably at least 95%, more preferably at least 98% and up to 99,6% identical to the human Prl-1, Prl-2, or Prl-3 variants, as defined in claim 2;
   (d) differs from the nucleic acid molecule of (a) to (c) by mutation and wherein said mutation causes an alteration, deletion, duplication or premature stop in the encoded polypeptide.

4. The composition of any one of claims 1-3, wherein the nucleic acid molecule is a DNA molecule, particularly a cDNA or a genomic DNA.

5. The composition of any one of claims 1-4, wherein said nucleic acid encodes a polypeptide contributing to regulating the energy homeostasis and/or the metabolism of triglycerides.

6. The composition of any one of claims 1-5, wherein said nucleic acid molecule is a recombinant nucleic acid molecule.

7. The composition of any one of claims 1-6, wherein the nucleic acid molecule is a vector, particularly an expression vector.

8. The composition of any one of claims 1-5, wherein the polypeptide is a recombinant polypeptide.

9. The composition of claim 8, wherein said recombinant polypeptide is a fusion polypeptide.

10. The composition of any one of claims 1-7, wherein said nucleic acid molecule is selected from hybridization probes, primers and anti-sense oligonucleotides.

11. The composition of any one of claims 1-10 which is a diagnostic composition.

12. The composition of any one of claims 1-10 which is a therapeutic composition.

13. The composition of any one of claims 1-12 for the manufacture of an agent for detecting and/or verifying, for the treatment, alleviation and/or prevention of metabolic diseases or dysfunctions, for example, but not limited to, metabolic syndrome including obesity, diabetes mellitus, eating disorder, cachexia, hypertension, coronary

heart disease, hypercholesterolemia (dyslipidemia), and gallstones, and others, in cells, cell masses, organs and/or subjects.

14. Use of a nucleic acid molecule of the protein tyrosine phosphatase gene family or a polypeptide encoded thereby or a fragment or a variant of said nucleic acid molecule or said polypeptide or an effector of said nucleic or polypeptide for controlling the function of a gene and/or a gene product which is influenced and/or modified by a protein tyrosine phosphatase polypeptide.

15. Use of the nucleic acid molecule of the protein tyrosine phosphatase gene family or a polypeptide encoded thereby or a fragment or a variant of said nucleic acid molecule or said polypeptide or an effector of said nucleic acid molecule or polypeptide for identifying substances capable of interacting with a protein tyrosine phosphatase polypeptide.

16. A non-human transgenic animal exhibiting a modified expression of a protein tyrosine phosphatase polypeptide.

17. The animal of claim 16, wherein the expression of a protein tyrosine phosphatase polypeptide is increased and/or reduced.

18. A recombinant host cell exhibiting a modified expression of a protein tyrosine phosphatase polypeptide.

19. The cell of claim 18 which is a human cell.

20. A method of identifying a (poly)peptide involved in the regulation of energy homeostasis and/or metabolism of triglycerides in a mammal comprising the steps of

(a) contacting a collection of (poly)peptides with a protein tyrosine phosphatase polypeptide or a fragment thereof under conditions that allow binding of said (poly)peptides;

(b) removing (poly)peptides which do not bind and

(c) identifying (poly)peptides that bind to said protein tyrosine phosphatase polypeptide.

21. A method of screening for an agent which modulates the interaction of a protein tyrosine phosphatase polypeptide with a binding target/agent, comprising the steps of

(a) incubating a mixture comprising

(aa) a protein tyrosine phosphatase polypeptide or a fragment thereof;

(ab) a binding target/agent of a protein tyrosine phosphatase polypeptide or fragment thereof; and

(ac) a candidate agent

under conditions whereby said Prl-1 polypeptide or fragment thereof specifically binds to said binding target/agent at a reference affinity;

(b) detecting the binding affinity of said Prl-1 polypeptide or fragment thereof to said binding target to determine an (candidate) agent-biased affinity; and

(c) determining a difference between (candidate) agent-biased affinity and the reference affinity.

22. A method of screening for an agent which modulates the activity of an protein tyrosine kinase polypeptide comprising the steps of

(a) incubating a mixture comprising

(aa) a protein tyrosine phosphatase polypeptide or a fragment thereof and

(ab) a candidate agent

under conditions whereby said Prl-1 polypeptide or fragment thereof exhibits a reference activity,

(b) detecting the activity of said Prl-1 polypeptide or fragment thereof to determine a (candidate) agent-biased activity and

(c) determining a difference between (candidate) agent-biased activity and reference activity.

23. A method of producing a composition comprising the (poly)peptide identified by the method of claim 20 or the agent identified by the method of claim 21 or 22 with a pharmaceutically acceptable carrier, diluent and/or adjuvant.

24. The method of claim 23 wherein said composition is a pharmaceutical composition for preventing, alleviating or treating of diseases and disorders, including metabolic diseases or dysfunctions, for example, but not limited to, such as metabolic syndrome including obesity, diabetes mellitus, eating disorder, cachexia, hypertension, coronary heart disease, hypercholesterolemia (dyslipidemia), and gallstones, and other diseases and disorders.

25. Use of a (poly)peptide as identified by the method of claim 20 or of an agent as identified by the method of claim 21 or 22 for the preparation of a pharmaceutical composition for the treatment, alleviation and/or prevention of of diseases and disorders, including metabolic diseases or dysfunctions, for example, but not limited to, metabolic syndrome including obesity, diabetes mellitus, eating disorder, cachexia, hypertension, coronary heart disease, hypercholesterolemia (dyslipidemia), and gallstones, and other diseases and disorders.

26. Use of a nucleic acid molecule of the protein tyrosine phosphatase gene family or of a fragment thereof for the preparation of a non-human animal which over- or under-expresses a protein tyrosine phosphatase product.

27. Kit comprising at least one of

(a) a protein tyrosine phosphatase nucleic acid molecule or a fragment thereof;

(b) a vector comprising the nucleic acid of (a);

(c) a host cell comprising the nucleic acid of (a) or the vector of (b);

(d) a polypeptide encoded by the nucleic acid of (a);

(e) a fusion polypeptide encoded by the nucleic acid of (a);

(f) an antibody, an aptamer or another receptor against the nucleic acid of (a) or the polypeptide of (d) or (e) and

(g) an anti-sense oligonucleotide of the nucleic acid of (a).

0 3. Dez. 2002

EPO - Munich
69
0 3. Dez. 2002

Abstract

The present invention discloses Prl-1 homologous proteins regulating the energy homeostasis and the metabolism of triglycerides, and polynucleotides, which identify and encode the proteins disclosed in this invention. The invention also relates to the use of these sequences in the diagnosis, study, prevention, and treatment of metabolic diseases and disorders.

ld 03.12.2002

0 3. Dez. 2002

EPO - Munich
69
0 3. Dez. 2002

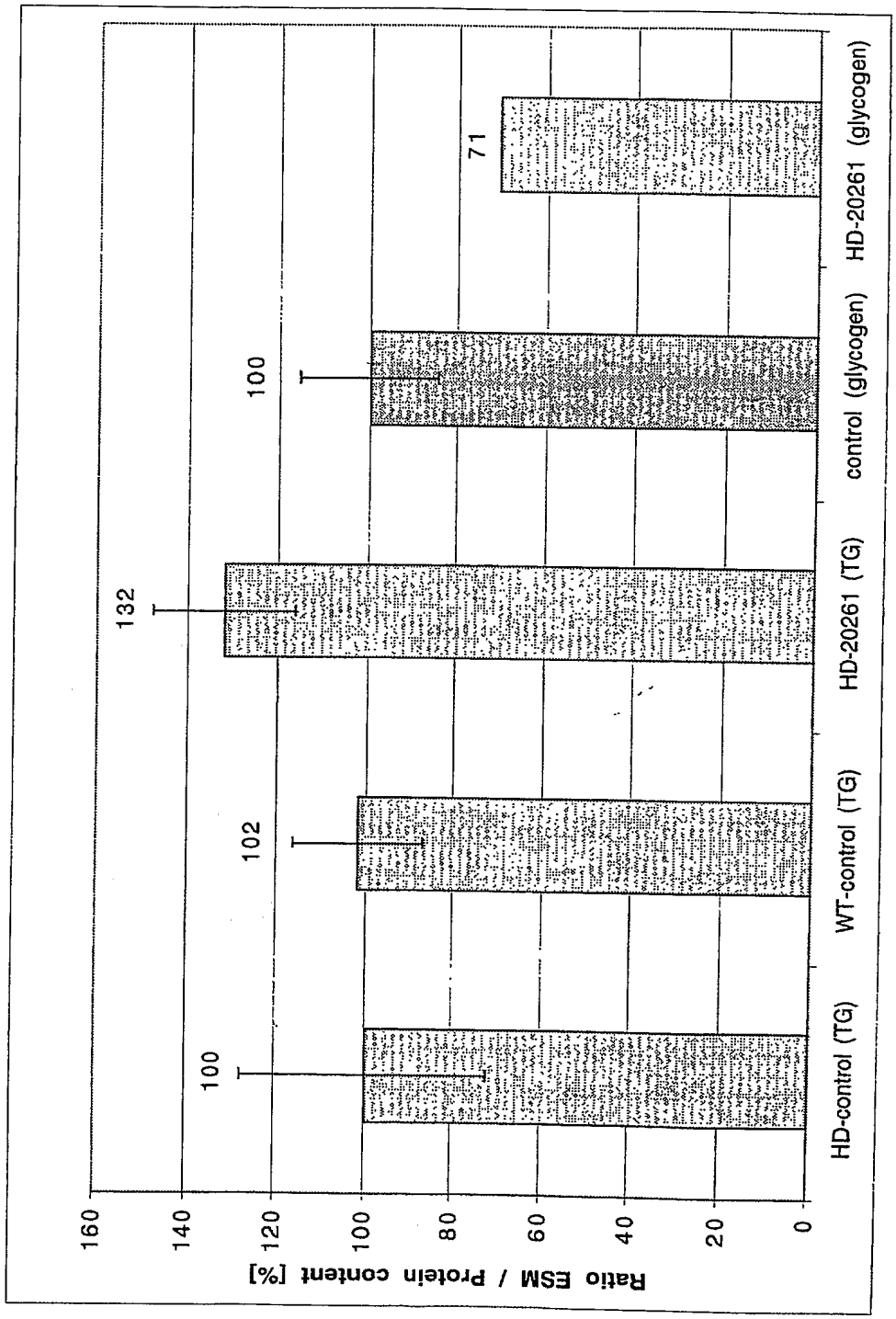

FIGURE 1. Energy storage metabolite content of a Drosophila PRL-1 mutant

Figure 2. Molecular organization of the PRL-1 gene (GadFly Acession Number CG4993)

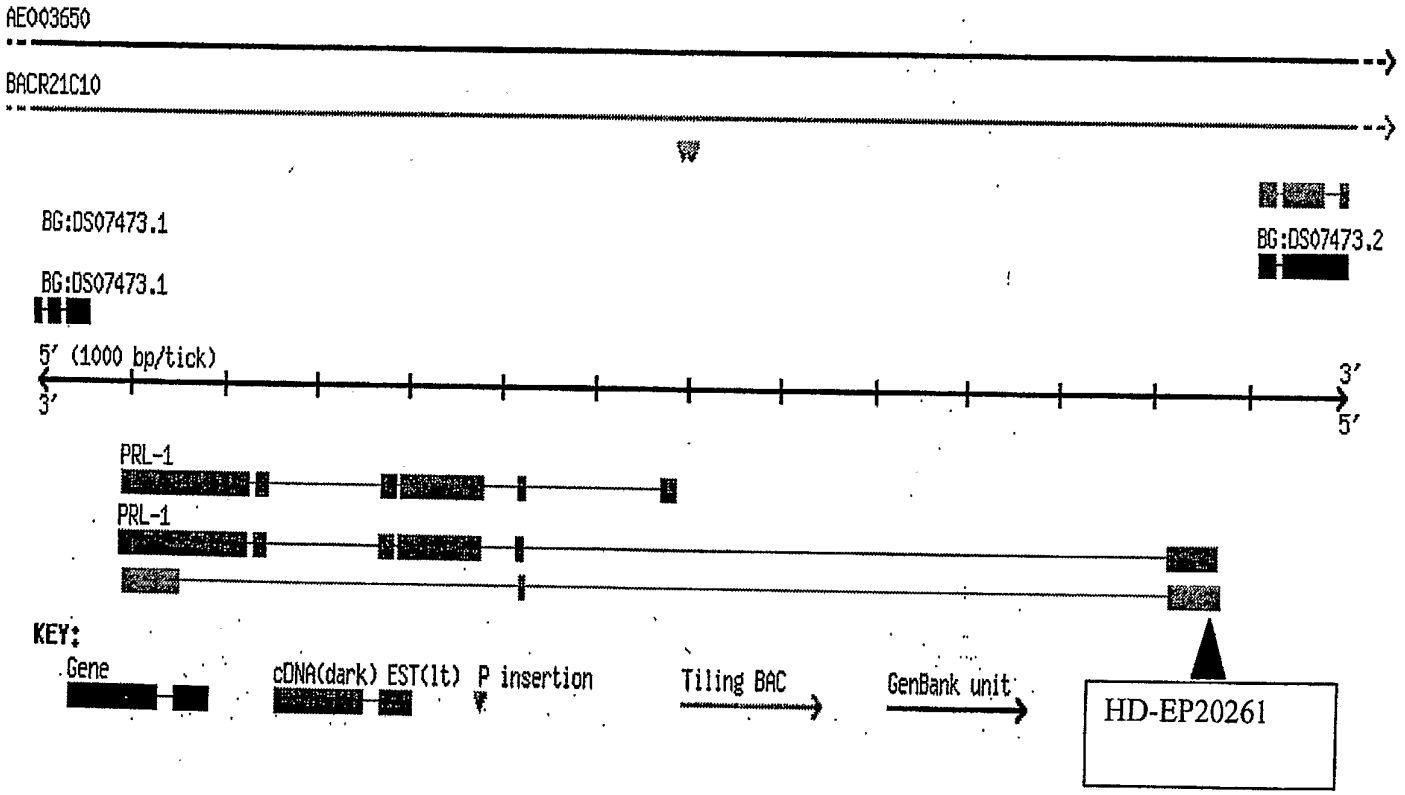

FIGURE 3. Nucleic acid sequences encoding the human proteins of the invention and amino acid sequences of the proteins of the invention

FIGURE 3A. Homo sapiens protein tyrosine phosphatase type IVA, member 1 (Prl-1), Nucleic acid sequence (SEQ ID NO: 1)

```
   1 ccggctcggt acgcgctctg ctccgagccg ctcactgcgt ggtagagtct ggtgcccccg
  61 ccgtcgcctg catcgccgcc accgccgctc cgccacgacc accgccgcct cctgccctgc
 121 agccaccgcc accgcctgtg tcgccgccgc ctcgggaccg gctgtatgat taggccacaa
 181 tcttcaatga gtaaacatat tcctcaattc tgtggtgttc ttggtcacac atttatggag
 241 tttctgaagg gcagtggaga ttactgccag gcacagcacg acctctatgc agacaagtga
 301 actgtagaaa ctgattactg ctccaccaag aagcccccat aagagtggtt atcctggaca
 361 cagaagtgtt gaattgaaat ccacagagca ttttacaaga gttctgacct ggatggggta
 421 aacctcagtg cacttctttt ctgttggcct cagtattact ggattgaaga attgctgctt
 481 cttgttagga ggttcatttc acttatcatt acttacaact tcatactcaa agcactgaga
 541 atttcaagtg gagtatattg aagtagactt cagtttcttt gcatcatttc tgtattcaat
 601 ttttttaatt atttcataac cctattgagt gtttttttaac taaattaaca tggctcgaat
 661 gaaccgccca gctcctgtgg aagtcacata caagaacatg agatttctta ttacacacaa
 721 tccaaccaat gcgaccttaa acaaatttat agaggaactt aagaagtatg gagttaccac
 781 aatagtaaga gtatgtgaag caacttatga cactactctt gtggagaaag aaggtatcca
 841 tgttcttgat tggccttttg atgatggtgc accaccatcc aaccagattg ttgatgactg
 901 gttaagtctt gtgaaaatta agtttcgtga agaacctggt tgttgtattg ctgttcattg
 961 cgttgcaggc cttgggagag ctccagtact tgttgcccta gcattaattg aaggtggaat
1021 gaaatacgaa gatgcagtac aattcataag acaaaagcgg cgtggagctt ttaacagcaa
1081 gcaacttctg tatttggaga agtatcgtcc taaaatgcgg ctgcgtttca aagattccaa
1141 cggtcataga aacaactgtt gcattcaata aaattggggt gcctaatgct actggaagtg
1201 gaacttgaga tagggcctaa tttgttatac atattagcca acatgttggc ttagtaagtc
1261 taatgaagct tccataggag tattgaaagg cagttttacc aggcctcaag ctagacagat
1321 ttggcaacct ctgtatttgg gttacagtca acctatttgg atacttggca aaagattctt
1381 gctgtcagca tataaaatgt gcttgtcatt tgtatcaatt gacctttccc caaatcatgc
1441 agtattgagt tatgacttgt taaatctatt cccatgccag aatcttatca atacataaga
1501 aatttaggaa gattaggtgc caaaataccc agcacaatac ttgtatattt ttagtaccat
1561 acagaagtaa aatcccagga actatgaaca ctagacctta tgtggtttat tccttcaatc
1621 atttcaaaca ttgaaagtag ggcctacatg gttatttgcc tgctcacttt atgtttacat
1681 ctcccacatt cataccaata tacgtcaggt ttgcttaacc attgattttt ttttttttt
1741 accaagtctt acagtgatta ttttacgtgt ttccatgtat ctcactttgt gctgtattaa
1801 aaaaacctcc attttgaaaa tctacgttgt acagaagcac atgtctttaa tgtcttcaga
1861 caaaaaagcc ttacattaat ttaatgtttg cactctgagg tgcaacttaa cagggagggc
1921 ctgagaaaag aatgggaggg ggctattaat tatttttag caaaatgttg cctttgtctt
1981 gtgcaaacat gtagaatatg ctctttaatt tagtaaaata tttttttaaa aggtagagat
2041 gctttgttat tgtaatcata aacttcctga aattcttgta attttttttcc catacttatc
2101 agaagtgtgt ttaccaactt atttttgttt gaaagtgtga ttttttttt ccttcccaac
2161 ctctcttgca aaaaaagaaa tgggtttctg ctaatgaatt gagcacatct aatattttat
2221 atgccttttg gagctgtgta agttaatatt tgatacttga caatttgttt tattatgtaa
2281 ttgataaaat ggtgatgtgt attaatgtta gttcaaccat atatttatac tgtctgggga
2341 tgtgtggtta tagttctgtg ggagaaataa ttttgtcagt gttcaccagc ttgtaaaaac
2401 ttagtgcgag agctgaaaca tctaaataaa taatgacatg catttatcat cattgagatt
2461 ggtttgcttt aaattaactt attttgtaga agacaaaatg aattgcactt cacttaatgt
2521 gtgtcctcat ctttttacaa ataaatgaag gattataaat gatgtcagca ttttagtaaa
2581 cttttagaca aaatttgtta gggtcattca tgaaaacttt aatactaaaa gcactttcca
2641 ttatatactt tttaaaggtc tagataattt tgaaccaatt tattattgtg tactgaggag
2701 aaataatgta tagtagagga cagccttggt ttgtaaagct cagttccact agttcatggt
2761 tttgtgcaac ttctgagcct cagttttctc ctttgcaaat taataattac atacctttat
2821 agattttgaa attaatttaa atattagtat ttggtacatg aaggcttaat gttaagtttc
2881 ctttaatgat ccacaataat ccctttgatc acgttaatct aaatctagat gtctttgtct
2941 aatttttttt gaatagcagt tataaatgta aaggactcaa agtttaagta aaaagtgata
```

```
3001 ctccaccttg tgtttcaaag aatttagttc cacctcttca taccagttta acacttaata
3061 tatttcattg gattttagac agggcaaaag gaagaacagg ggcctctgga ggcccttggt
3121 tatttaaatc ttggattatt tgtgatagta atcacaaatt tttggctaat ttttaacctg
3181 aggttttgtt tttttttaa aggaaatgca gcctagtctt gagaacataa ttttatataa
3241 tcaattacta aatgttaaac tattaccaca cagcccataa aacagcattt gcgtttattg
3301 agagagagga tgtgccatca tgattaatga aaactatctt ttgagtttga aaagaaatta
3361 atttgcagtg tttggattgt atatatggtg ctaaaaataa attaatttac tttataaacc
3421 ttatctgtac attatacgat gtgatgaaat ttgcttttta tccaaatatt ttgtatcttg
3481 taaatatggc taattatagg aatgcctata atacatctta gattccttat atctaataag
3541 agttcaaaga gttatgagtt gaagtcttga atgcaggaaa ctatctgata gtgttctaaa
3601 atttggttac ttgggtttgg ataccccttag tgggatgatg taaatagagg ctagctacct
3661 aggcttgtct atagcaacca taatgttgat gtaagtaatg cggttactga atcataagaa
3721 aatgccatct ctttttagtt gaaggaaaac tctggaagta ggtgccattg gtcattctgc
3781 agtgcactgc aaccattgtt tcccctagtg ccctcttttc cctagggcat tgctctccta
3841 ttcccacgcc ttaacacagc tctataccta gaagcagcca gcccaggcat gcagtcacat
3901 ttaatcacat ccccccttcta gagtgcttca aaatgatgta gtccctcaac ttggctaaag
3961 aatctcaatc tcttgaaatt tatttttta atgtcatatt catctggtaa atatctactg
4021 tttgccaggc atttaagaat atggcaaaga acataaaaga tggtgtcacc agattttggt
4081 caccaatgag tacccgaccc gttgccatga ttaagagaga atgctttcta ttggagtttc
4141 aggaaatata atttgagaat actttaaagg gaagtggaag tataagtgaa tgatattttt
4201 cttttacatg taaacaatga agttatttca aagttaagtt ttaaacaaaa tacatgaagt
4261 agtgtctgcc atacatgtta atattctaca ttcttgcttc cttaaattaa tatgtttgtg
4321 tgtatatatg tgcctcacac ctgaattgaa aattaaagac tggtttaaaa gtggttaaaa
4381 aaaaaaaaaa aaaa
```

FIGURE 3B. Homo sapiens protein tyrosine phosphatase type IVA, member 1 (Prl-1), Amino acid sequence (SEQ ID NO: 2)

```
  1 marmnrpapv evtyknmrfl ithnptnatl nkfieelkky gvttivrvce atydttlvek
 61 egihvldwpf ddgappsnqi vddwlslvki kfreepgcci avhcvaglgr apvlvalali
121 eggmkyedav qfirqkrrga fnskqllyle kyrpkmrlrf kdsnghrnnc ciq
```

FIGURE 3C. Homo sapiens protein tyrosine phosphatase type IVA, member 2 (Prl-2), Nucleic acid sequence, transcript variant 1 (SEQ ID NO: 3)

```
   1 agcggggctg cgcgaagtca tcgctgttcc agacagcgat gactcgagag cggtgggggt
  61 ggcggcgcga tcggccgggc tgtaaccgtc gtctgtccgg gagcggctgg agcggcagcg
 121 gcggccgggc acggcgcgag gtgacgccac agggcagcgg cggcagcgga ggcagcggcg
 181 gcagcaggag acgcagcggc ggccgcagca gcagcagcaa gacggactcg tggagacgcg
 241 ccgccgccgc cgccgccggg ccgggccggg tgtcgcgcgc cgaggctggg ggggagtcgt
 301 cgccgccgcc gccaccgcta ccgccgccgc cgccgccgcc gaggtgactg aggagagagg
 361 cgcctcctcg ctcccgccac cgccggactt caatgcccag tccccagctc gccagcgttt
 421 ttcgttggaa tatacgttgc acatttatgg cgattctgag tgtgagggca gacttctgcc
 481 aggctcagca cagcattttc gctgacaagt gagcttggag gttctatgtg ccataattaa
 541 cattgccttg aagactcctg gacaccgaga ctggcctcag aaatagttgg ctttttttt
 601 tttttaattg caagcatatt tcttttaatg actccagtaa aattaagcat caagtaaaca
 661 agtggaaagt gacctacact tttaacttgt ctcactagtg cctaaatgta gtaaaggctg
 721 cttaagtttt gtatgtagtt ggatttttg gagtccgaat atttccatct gcagaaattg
 781 aggcccaaat tgaatttgga ttcaagtgga ttctaaatac tttgcttatc ttgaagagag
 841 aagcttcata aggaataaac aagttgaata gagaaaacac tgattgataa taggcatttt
 901 agtggtcttt ttaatgtttt ctgctgtgaa acatttcaag atttattgat ttttttttt
 961 cactttcccc atcacactca cacgcacgct cacacttttt atttgccata atgaaccgtc
1021 cagcccctgt ggagatctcc tatgagaaca tgcgttttct gataactcac aaccctacca
1081 atgctactct caacaagttc acagaggaac ttaagaagta tggagtgacg actttggttc
1141 gagtttgtga tgctacatat gataaagctc cagttgaaaa agaaggaatc cacgttctag
1201 attggccatt tgatgatgga gctccacccc ctaatcagat agtagatgat tggttaaacc
```

```
1261 tgttaaaaac caaatttcgt gaagagccag gttgctgtgt tgcagtgcat tgtgttgcag
1321 gattgggaag ggcacctgtg ctggttgcac ttgctttgat tgaatgtgga atgaagtacg
1381 aagatgcagt tcagtttata agacaaaaaa gaaggggagc gttcaattcc aaacagctgc
1441 tttatttgga gaaataccga cctaagatgc gattacgctt cagagatacc aatgggcatt
1501 gctgtgttca gtagaaggaa atgtaaacga aggctgactt gattgtgcca tttagaggga
1561 actcttggta cctggaaatg tgaatctgga atattacctg tgtcatcaaa gtagtgatgg
1621 attcagtact cctcaaccac tctcctaatg attggaacaa aagcaaacaa aaaagaaatc
1681 tctctataaa atgaataaaa tgtttaagaa aagagaaaga gaaaaggaat taattcagtg
1741 aaggatgatt ttgctcctag ttttggagtt tgaatttctg ccaggattga attattttga
1801 aatctcctgt ctttttaaac tttttcaaaa taggtctcta aggaaaacca gcagaacatt
1861 aggcctgtgc aaaaccatct gtttggggag cacactcttc cattatgctt ggcacataga
1921 tctccctgtg gtgggatttt ttttttccct ttttttgtgg gggagggttg gtggtatatt
1981 tttcccctct tttttccttc ctctcctaca tctccctttt cccccgatcc aagttgtaga
2041 tggaatagaa gcccttgttg ctgtagatgt gcgtgcagtc tggcagcctt aagcccacct
2101 gggcactttt agataaaaaa aaaaaaaaac aaaaaacaac accaaaaaaa cagcagtgat
2161 atatatattc caggtggttt ttagtcttta ctgatgaaag ggtgttcatg ttagtttctt
2221 caaaacccta tctaatacta ggcaaagtag ccaagagcct tttgttttgt ttttatttttg
2281 ataaattagt ggagaaatgg cattttaaga ggagtctctt ctcaacttac ctgagagtcg
2341 aattcttctc ttccctaacc aatgaagcta agtggttatc ccagaaactt gtcttctaaa
2401 agggaggact ccaggccatc aataaagatg tccaggcagt gagcgtactt tttacaccct
2461 gtagaattgt gggctgtagc gttactctga ttttctgtct agtatcagag aatgctggta
2521 gcttaaaatt tttattttag gacttgtact ctgaattttc aggaaccgtc aaaggagcag
2581 cagcaaattc acatattttc gacttgagaa atgcttgtgg tatgtgtttt ccaaactgcc
2641 ccctatatgt aaagttcagt ttaaccactg attgccttgt tattactagg ttttttgaga
2701 ttaaaaaaaa aaaatccctg gtttaaaacc aacaatgatg cctagtgagt atgtgtccac
2761 aggccataac agggtagaag agagacatcg tgcaacccaa tgagtagtga agggactgtg
2821 ttgcttgtga agcggtgtag tagcattttt gcagattctt ggctgggttt agtgtactga
2881 tctagaaaag ctgtttttct gctcctttgt ggaaggcagt tatgatcagg ctgcatggac
2941 aaagcaggta gaggggcacc atcaggggct cttgcactat tttcacctct aaatattacg
3001 tactcagtag tgccctgctt ctagggctct gaatacgggc ttaaagtcat cttgtcctgc
3061 tggaatttgc tgtgcagagc cataagcctc ccattttgtt agcgtcagct aggccaatag
3121 gaacagaccg ggaccttgtc tcacactgat gatacctcac atgttgaccg gctatgtgaa
3181 ctgcctattt cctatgctgg agttttgatt tttaactaaa cgcaaatctg tagattctct
3241 cctctcccat cccagaaaac aaaacaaaat aatgcttttc gaaattgttt ctaggacttt
3301 aaaacataat ggtatatcca aaattcttta tttcagaatg caacaataga ttccattaat
3361 atagactcaa gatcaaaaca gcatacctgc taagctaaga tagatggtgt tgattccact
3421 gggttttgat caatacaata acaaaccttt ttcctttgac atactctgaa ttttgttgtt
3481 tgggggggagg gggtgtgtgt gtgtgtgtgt gtgtgtgtgt gtattgtgtg tgtgtgtgtg
3541 tgcacgcgca gtgtccatca gtatcagtgc ctgcctgagt taggaaaatt acattcctgg
3601 ttctgtattg aggagaagga tgtataaagc aacatgaaac attagccctc cttttatttt
3661 aaagactaat gttaattgtt cttaaaactg gatttttttt ccttaaagca atttttttct
3721 tttcgattta atgaagtatt gctagctgaa gccagtttga catagagaga tgtcagattg
3781 atttgaaagg tgtgcagcct gatttaaaac caaaccctga acccttttaa agaacaataa
3841 aacatatttt acacgctcaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa
3901 aaaaaaaaaa aaaaaaaaaa aaaaa
```

FIGURE 3D. Homo sapiens protein tyrosine phosphatase type IVA, member 2 (Prl-2), Amino acid sequence transcript variant 1 (SEQ ID NO: 4)

```
  1 mnrpapveis yenmrflith nptnatlnkf teelkkygvt tlvrvcdaty dkapvekegi
 61 hvldwpfddg apppnqivdd wlnllktkfr eepgccvavh cvaglgrapv lvalaliecg
121 mkyedavqfi rqkrrgafns kqllylekyr pkmrlrfrdt nghccvq
```

FIGURE 3E. Homo sapiens protein tyrosine phosphatase type IVA, member 2 (Prl-2), Nucleic acid sequence, transcript variant 2 (SEQ ID NO: 5)

```
   1 agcggggctg cgcgaagtca tcgctgttcc agacagcgat gactcgagag cggtgggggt
  61 ggcggcgcga tcggccgggc tgtaaccgtc gtctgtccgg gagcggctgg agcggcagcg
 121 gcggccgggc acggcgcgag gtgacgccac agggcagcgg cggcagcgga ggcagcggcg
 181 gcagcaggag acgcagcggc ggccgcagca gcagcagcaa gacggactcg tggagacgcg
 241 ccgccgccgc cgccgccggg ccgggccggg tgtcgcgcgc cgaggctggg ggggagtcgt
 301 cgccgccgcc gccaccgcta ccgccgccgc cgccgccgcc gaggtgactg aggagagagg
 361 cgcctcctcg ctcccgccac cgccggactt caatgcccag tccccagctc gccagcgttt
 421 ttcgttggaa tatacgttgc acatttatgg cgattctgag tgtgagggca gacttctgcc
 481 aggctcagca cagcattttc gctgacaagt gagcttggag gttctatgtg ccataattaa
 541 cattgccttg aagactcctg gacaccgaga ctggcctcag aaatagttgg ctttttttt
 601 tttttaattg caagcatatt tcttttaatg actccagtaa aattaagcat caagtaaaca
 661 agtggaaagt gacctacact tttaacttgt ctcactagtg cctaaatgta gtaaaggctg
 721 cttaagtttt gtatgtagtt ggatttttg gagtccgaat atttccatct gcagaaattg
 781 aggcccaaat tgaatttgga ttcaagtgga ttctaaatac tttgcttatc ttgaagagag
 841 aagcttcata aggaataaac aagttgaata gagaaaacac tgattgataa taggcatttt
 901 agtggtcttt ttaatgtttt ctgctgtgaa acatttcaag atttattgat ttttttttt
 961 cactttcccc atcacactca cacgcacgct cacacttttt atttgccata atgaaccgtc
1021 cagcccctgt ggagatctcc tatgagaaca tgcgttttct gataactcac aaccctacca
1081 atgctactct caacaagttc acagaggaac ttaagaagta tggagtgacg actttggttc
1141 gagtttgtga tgctacatat gataaagctc cagttgaaaa agaaggaatc cacgttctag
1201 attggccatt tgatgatgga gctccacccc ctaatcagat agtagatgat tggttaaacc
1261 tgttaaaaac caaatttcgt gaagagccag gttgctgtgt tgcagtgcat tgtgttgcag
1321 gattgggaag ggcacctgtg ctggttgcac ttgctttgat tgaatgtgga atgaagtacg
1381 aagatgcagt tcagtttata agacaaaaaa gaaggggagc gttcaattcc aaacagctgc
1441 tttatttgga gaaataccga cctaagatgc gattacgctt cagagatacc aatgggcatt
1501 gctgtgttca gtagaaggaa atgtaaacga aggctgactt gattgtgcca tttagaggga
1561 actcttggta cctggaaatg tgaatctgga atattacctg tgtcatcaaa gtagtgatgg
1621 attcagtact cctcaaccac tctcctaatg attggaacaa aagcaaacaa aaaagaaatc
1681 tctctataaa atgaataaaa tgtttaagaa aagagaaaga gaaaaggaat taattcagtg
1741 aaggatgatt ttgctcctag ttttggagtt tgaatttctg ccaggattga attattttga
1801 aatctcctgt ctttttaaac tttttcaaaa taggtctcta aggaaaacca gcagaacatt
1861 aggcctgtgc aaaaccatct gtttgggggag cacactcttg gcacatagat ctccctgtgg
1921 tgggattttt ttccctttt ttgtgggga gggttggtgg tatatttttc ccctcttttt
1981 tccttcctct cctacatctc ccttttcccc cgatccaagt tgtagatgga atagaagccc
2041 ttgttgctgt agatgtgcgt gcagtctggc agccttaagc ccac
```

FIGURE 3F. Homo sapiens protein tyrosine phosphatase type IVA, member 2 (Prl-2), Amino acid sequence transcript variant 2 (SEQ ID NO: 6)

```
  1 mnrpapveis yenmrflith nptnatlnkf teelkkygvt tlvrvcdaty dkapvekegi
 61 hvldwpfddg apppnqivdd wlnllktkfr eepgccvavh cvaglgrapv lvalaliecg
121 mkyedavqfi rqkrrgafns kqllylekyr pkmrlrfrdt nghccvq
```

FIGURE 3G. Homo sapiens protein tyrosine phosphatase type IVA, member 2 (Prl-2), Nucleic acid sequence, transcript variant 3 (SEQ ID NO: 7)

```
   1 agcggggctg cgcgaagtca tcgctgttcc agacagcgat gactcgagag cggtgggggt
  61 ggcggcgcga tcggccgggc tgtaaccgtc gtctgtccgg gagcggctgg agcggcagcg
 121 gcggccgggc acggcgcgag gtgacgccac agggcagcgg cggcagcgga ggcagcggcg
 181 gcagcaggag acgcagcggc ggccgcagca gcagcagcaa gacggactcg tggagacgcg
 241 ccgccgccgc cgccgccggg ccgggccggg tgtcgcgcgc cgaggctggg ggggagtcgt
 301 cgccgccgcc gccaccgcta ccgccgccgc cgccgccgcc gaggtgactg aggagagagg
 361 cgcctcctcg ctcccgccac cgccggactt caatgcccag tccccagctc gccagcgttt
 421 ttcgttggaa tatacgttgc acatttatgg cgattctgag tgtgagggca gacttctgcc
 481 aggctcagca cagcattttc gctgacaagt gagcttggag gttctatgtg ccataattaa
 541 cattgccttg aagactcctg gacaccgaga ctggcctcag aaatagttgg ctttttttt
 601 tttttaattg caagcatatt tcttttaatg actccagtaa aattaagcat caagtaaaca
 661 agtggaaagt gacctacact tttaacttgt ctcactagtg cctaaatgta gtaaaggctg
 721 cttaagtttt gtatgtagtt ggatttttg gagtccgaat atttccatct gcagaaattg
 781 aggcccaaat tgaatttgga ttcaagtgga ttctaaatac tttgcttatc ttgaagagag
 841 aagcttcata aggaataaac aagttgaata gagaaaacac tgattgataa taggcatttt
 901 agtggtcttt ttaatgtttt ctgctgtgaa acatttcaag atttattgat ttttttttt
 961 cactttcccc atcacactca cacgcacgct cacacttttt atttgccata atgaaccgtc
1021 cagcccctgt ggagatctcc tatgagaaca tgcgttttct gataactcac aaccctacca
1081 atgctactct caacaagttc acagaggaac ttaagaagta tggagtgacg actttggttc
1141 gagtttgtga tgctacatat gataaagctc cagttgaaaa agaaggaatc cacgttctaa
1201 aaaagaaggg gagcgttcaa ttccaaacag ctgctttatt tggagaaata ccgacctaag
1261 atgcgattac gcttcagaga taccaatggg cattgctgtg ttcagtagaa ggaaatgtaa
1321 acgaaggctg acttgattgt gccatttaga gggaactctt ggtacctgga aatgtgaatc
1381 tggaatatta cctgtgtcat caaagtagtg atggattcag tactcctcaa ccactctcct
1441 aatgattgga acaaaagcaa acaaaaaaga aatctctcta taaaatgaat aaaatgttta
1501 agaaaagaga aagagaaaag gaattaattc agtgaaggat gattttgctc ctagttttgg
1561 agtttgaatt tctgccagga ttgaattatt ttgaaatctc ctgtcttttt aaacttttc
1621 aaaataggtc tctaaggaaa accagcagaa cattaggcct gtgcaaaacc atctgtttgg
1681 ggagcacact ctt
```

FIGURE 3H. Homo sapiens protein tyrosine phosphatase type IVA, member 2 (Prl-2), Amino acid sequence, transcript variant 3 (SEQ ID NO: 8)

```
 1 mnrpapveis yenmrflith nptnatlnkf teelkkygvt tlvrvcdaty dkapvekegi
61 hvlkkkgsvq fqtaalfgei pt
```

FIGURE 3I. Homo sapiens protein tyrosine phosphatase type IVA, member 3 (Prl-3), Nucleic acid sequence (SEQ ID NO: 9)

```
   1 tgactatcca gctctgagag acgggagttt ggagttgccc gctttacttt ggttgggttg
  61 gggggggcgg cgggctgttt tgttcctttt ctttttaag agttgggttt tctttttaa
 121 ttatccaaac agtgggcagc ttcctccccc acacccaagt atttgcacaa tatttgtgcg
 181 gggtatgggg gtgggttttt aaatctcgtt tctcttggac aagcacaggg atctcgttct
 241 cctcattttt tgggggtgtg tggggacttc tcaggtcgtg tccccagcct tctctgcagt
 301 cccttctgcc ctgccgggcc cgtcgggagg cgccatggct cggatgaacc gcccggcccc
 361 ggtggaggtg agctacaaac acatgcgctt cctcatcacc cacaacccca ccaacgccac
 421 gctcagcacc ttcattgagg acctgaagaa gtacgggggct accactgtgg tgcgtgtgtg
 481 tgaagtgacc tatgacaaaa cgccgctgga gaaggatggc atcaccgttg tggactggcc
 541 gtttgacgat ggggcgcccc cgcccggcaa ggtagtggaa gactggctga gcctggtgaa
 601 ggccaagttc tgtgaggccc ccggcagctg cgtggctgtg cactgcgtgg cgggcctggg
```

```
 661 ccgggctcca gtccttgtgg cgctggcgct tattgagagc gggatgaagt acgaggacgc
 721 catccagttc atccgccaga agcgccgcgg agccatcaac agcaagcagc tcacctacct
 781 ggagaaatac cggcccaaac agaggctgcg gttcaaagac ccacacacgc acaagacccg
 841 gtgctgcgtt atgtagctca ggaccttggc tgggcctggt cgtcatgtag gtcaggacct
 901 tggctggacc tggaggccct gcccagccct gctctgccca gcccagcagg ggctccaggc
 961 cttggctggc cccacatcgc cttttcctcc ccgacacctc cgtgcacttg tgtccgagga
1021 gcgaggagcc cctcgggccc tgggtggcct ctgggccctt tctcctgtct ccgccactcc
1081 ctctggcggc gctggccgtg gctctgtctc tctgaggtgg gtcgggcgcc ctctgcccgc
1141 cccctcccac accagccagg ctggtctcct ctagcctgtt tgttgtgggg tgggggtata
1201 ttttgtaacc actgggcccc cagcccctct tttgcgaccc cttgtcctga cctgttctcg
1261 gcaccttaaa ttattagacc ccggggcagt caggtgctcc ggacacccga aggcaataaa
1321 acaggagccg tgaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa
1381 aaaaaaaaaa aaaaaa
```

FIGURE 3J. Homo sapiens protein tyrosine phosphatase type IVA, member 3 (Prl-3), Amino acid sequence (SEQ ID NO: 10)

```
  1 marmnrpapv evsykhmrfl ithnptnatl stfiedlkky gattvvrvce vtydktplek
 61 dgitvvdwpf ddgapppgkv vedwlslvka kfceapgscv avhcvaglgr apvlvalali
121 esgmkyedai qfirqkrrga inskqltyle kyrpkqrlrf kdphthktrc cvm
```

Figure 4. Expression of Prl-1 in different mammalian models

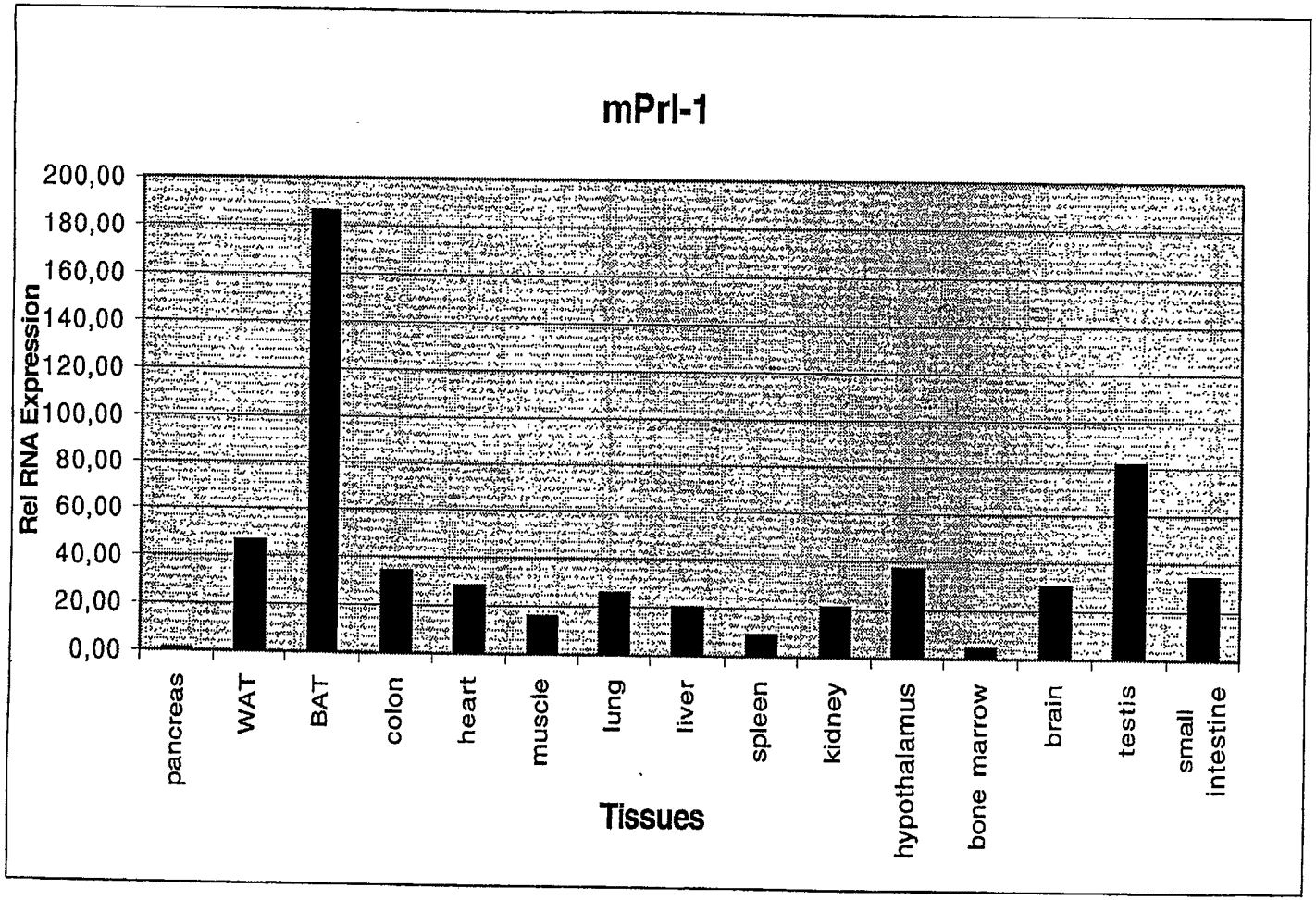

Figure 4A. Real-time PCR analysis of Prl-1 expression in wild type mouse tissues

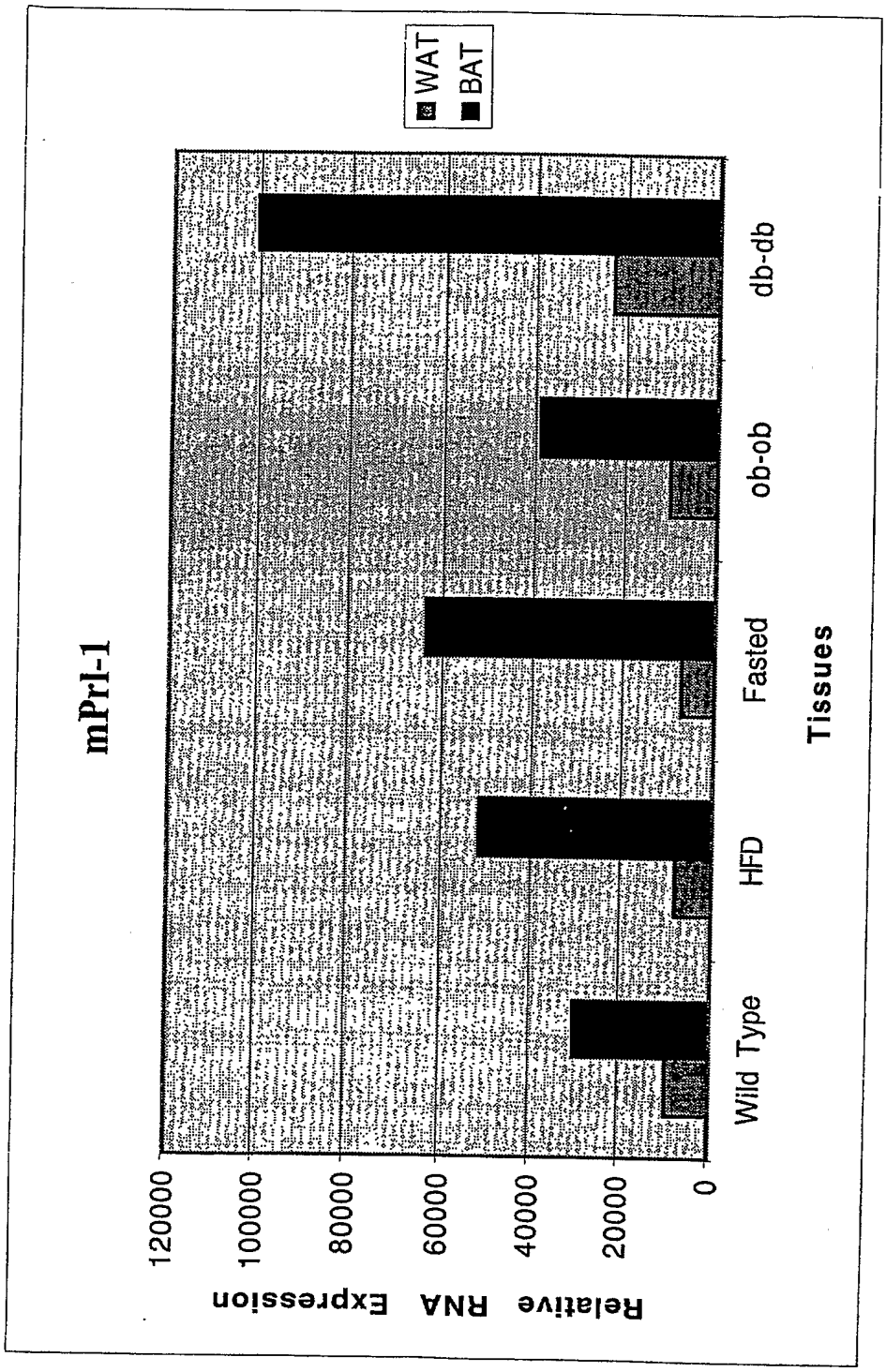

Figure 4B. Prl-1 expression in metabolically active tissues

FIGURE 5. In vitro assays for the determination of triglyceride storage and glycogen levels in adipocytes overexpressing Prl-1

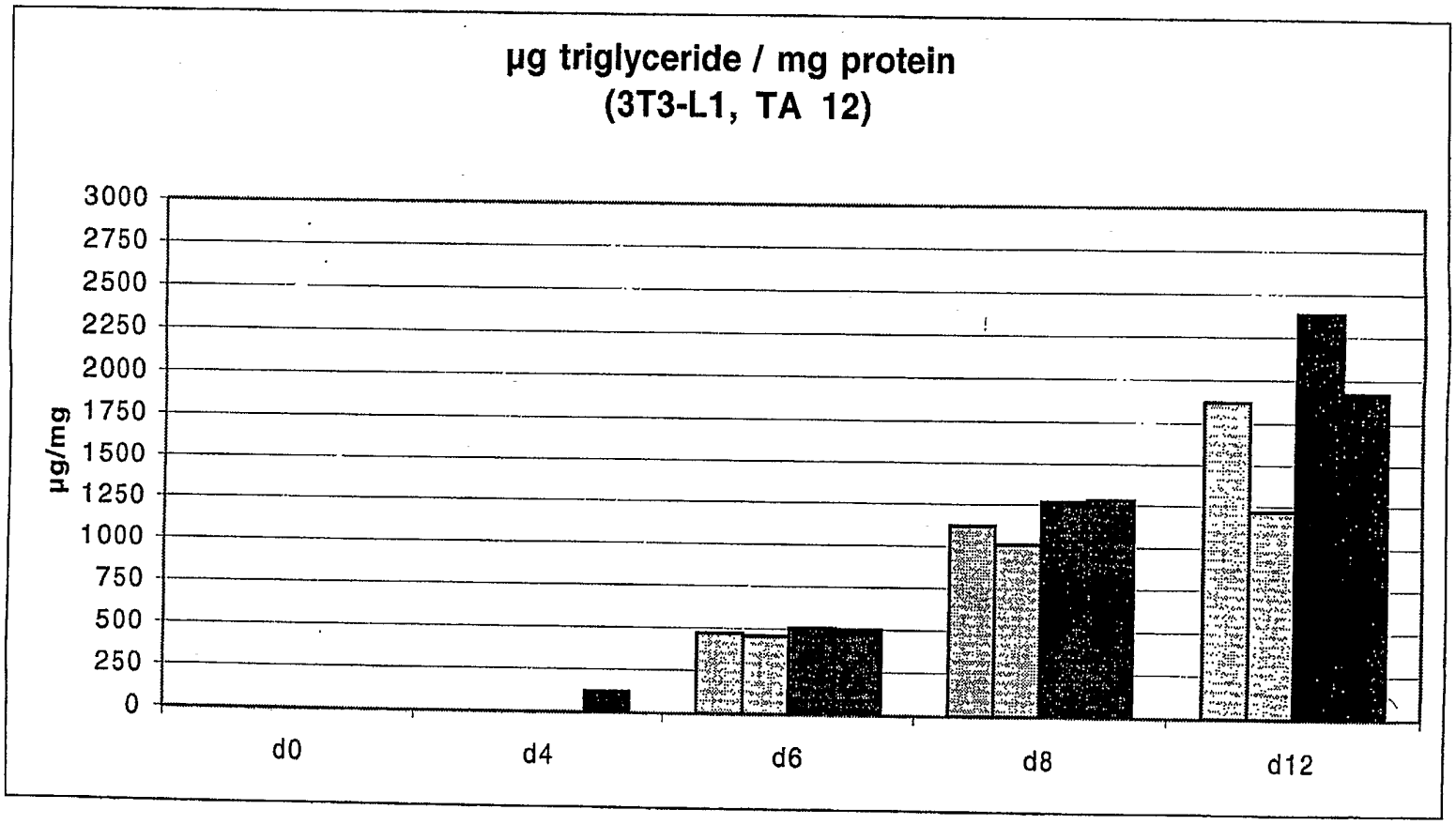

Figure 5A shows upregulation of cellular triglyceride levels in cells overexpressing Prl-1

11/12

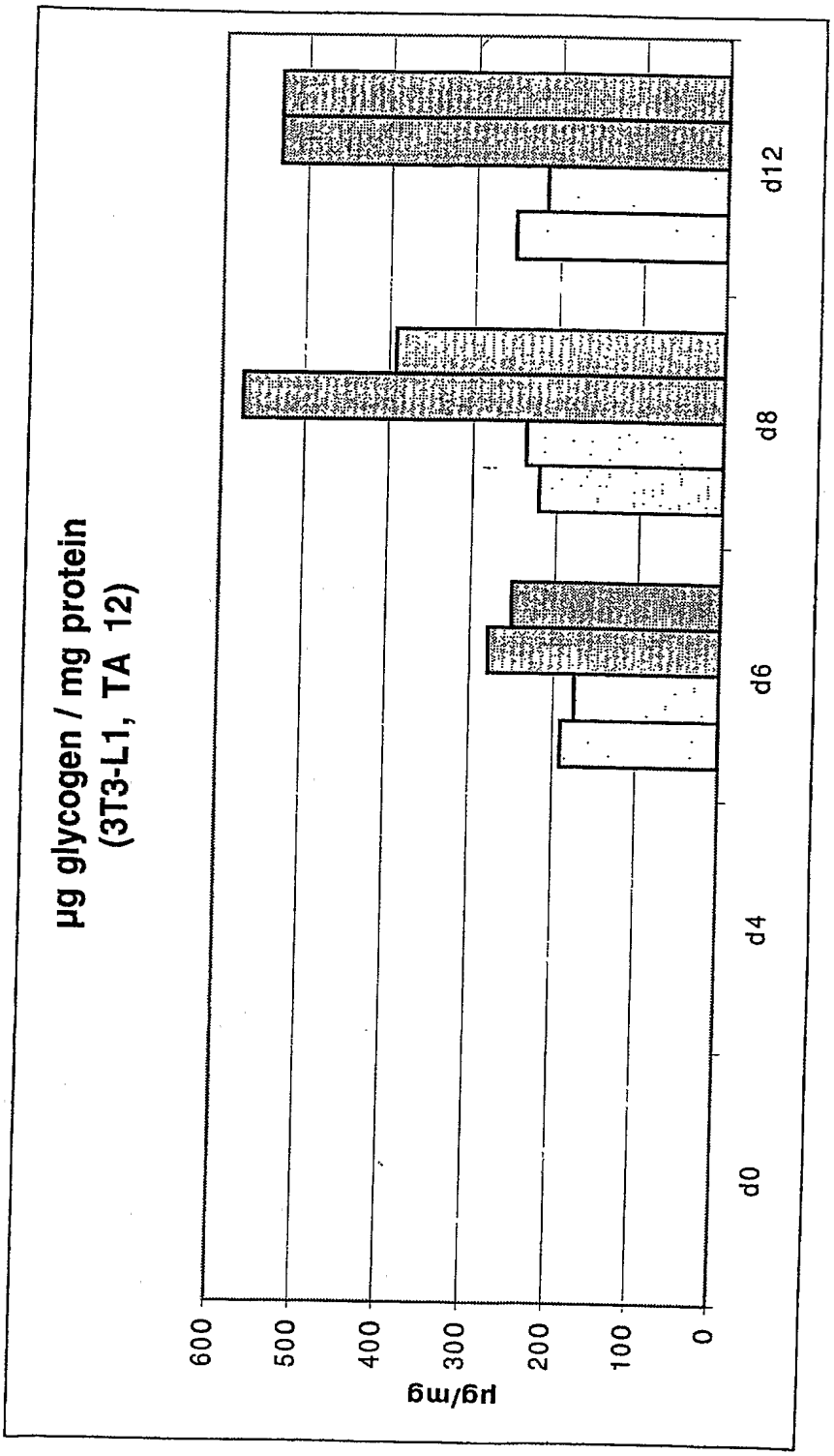

Figure 5B shows the upregulation of cellular glycogen levels in cells overexpressing Prl-1